

NONAUTONOMOUS KOLMOGOROV PARABOLIC EQUATIONS WITH UNBOUNDED COEFFICIENTS

MARKUS KUNZE, LUCA LORENZI, AND ALESSANDRA LUNARDI

ABSTRACT. We study a class of elliptic operators A with unbounded coefficients defined in $I \times \mathbb{R}^d$ for some unbounded interval $I \subset \mathbb{R}$. We prove that, for any $s \in I$, the Cauchy problem $u(s, \cdot) = f \in C_b(\mathbb{R}^d)$ for the parabolic equation $D_t u = Au$ admits a unique bounded classical solution u . This allows to associate an evolution family $\{G(t, s)\}$ with A , in a natural way. We study the main properties of this evolution family and prove gradient estimates for the function $G(t, s)f$. Under suitable assumptions, we show that there exists an evolution system of measures for $\{G(t, s)\}$ and we study the first properties of the extension of $G(t, s)$ to the L^p -spaces with respect to such measures.

1. INTRODUCTION AND SUMMARY

Parabolic partial differential equations with unbounded coefficients occur naturally in the study of stochastic processes. Let us consider the stochastic differential equation

$$(1.1) \quad \begin{cases} dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, & t > s, \\ X_s = x. \end{cases}$$

Here, W_t is a standard d -dimensional Brownian motion and μ (resp. σ) are regular \mathbb{R}^d (resp. $\mathbb{R}^{d \times d}$) valued coefficients. If (1.1) has a solution $X_t = X(t, s, x)$ for all $x \in \mathbb{R}^d$, it follows from Itô's formula that, for $\varphi \in C_b^2(\mathbb{R}^d)$ and $t \in \mathbb{R}$, the function

$$u(s, x) := \mathbb{E}(\varphi(X(t, s, x)))$$

solves the partial differential equation

$$(1.2) \quad \begin{cases} u_s(s, x) = -\frac{1}{2}\text{Tr}((\sigma(s, x)\sigma^*(s, x))D_x^2u(s, x)) - \langle \mu(s, x), \nabla_x u(s, x) \rangle, & s < t, \\ u(t, x) = \varphi(x). \end{cases}$$

This shows how probability theory may be used to obtain information about the solutions of second order evolution PDE's. In the case of Lipschitz continuous coefficients, there are many results stating conditions on μ and σ such that (1.1) is well posed. See, e.g., [12, 13, 14].

It is also possible to take (1.2) as a starting point and work in a purely analytic manner. This has been done in several papers in the autonomous case (see e.g., the book [3] and its bibliography). To the best of our knowledge, in the literature there is not any systematic treatment of the nonautonomous case except in the particular case when the elliptic operator in (1.2) is the non autonomous Ornstein-Uhlenbeck operator (see [6, 10, 11]).

Work supported by the M.I.U.R. research projects Prin 2004 and 2006 “Kolmogorov equations”.
To appear on Trans. Amer. Math. Soc.

In this paper we set the basis for the general theory of non autonomous operators. More precisely, we consider the equation

$$(1.3) \quad \begin{cases} u_t(t, x) = \mathcal{A}(t)u(t, x), & (t, x) \in (s, +\infty) \times \mathbb{R}^d, \\ u(s, x) = f(x), & x \in \mathbb{R}^d. \end{cases}$$

The operators $\mathcal{A}(t)$ appearing in (1.2) are defined on smooth functions φ by

$$\begin{aligned} (\mathcal{A}(t)\varphi)(x) &= \sum_{i,j=1}^d q_{ij}(t, x)D_{ij}\varphi(x) + \sum_{i=1}^d b_i(t, x)D_i\varphi(x) \\ &= \text{Tr}(Q(t, x)D^2\varphi(x)) + \langle b(t, x), \nabla\varphi(x) \rangle. \end{aligned}$$

The time index t varies over an interval I which is either \mathbb{R} or a right halfline.

Note that the equation in (1.3) is forward in time in contrast to equation (1.2). However, reverting time, solutions of (1.3) are transformed into solutions of (1.2) and viceversa. Our standing hypotheses on the data $b = (b_i)$ and $Q = (q_{ij})$ are the following:

- Hypothesis 1.1.**
- (i) The coefficients q_{ij} and b_i belong to $C_{\text{loc}}^{\frac{\alpha}{2}, \alpha}(I \times \mathbb{R}^d)$ for any $i, j = 1, \dots, d$ and some $\alpha \in (0, 1)$;
 - (ii) Q is uniformly elliptic, i.e., for every $(t, x) \in I \times \mathbb{R}^d$, the matrix $Q(t, x)$ is symmetric and there exists a function $\eta : I \times \mathbb{R}^d \rightarrow \mathbb{R}$ such that $0 < \eta_0 := \inf_{I \times \mathbb{R}^d} \eta$ and

$$\langle Q(t, x)\xi, \xi \rangle \geq \eta(t, x)|\xi|^2, \quad \xi \in \mathbb{R}^d, \quad (t, x) \in I \times \mathbb{R}^d;$$

- (iii) for every bounded interval $J \subset I$ there exist a function $\varphi = \varphi_J \in C^2(\mathbb{R}^d)$ and a positive number $\lambda = \lambda_J$ such that

$$\lim_{|x| \rightarrow +\infty} \varphi(x) = +\infty \quad \text{and} \quad (\mathcal{A}(t)\varphi)(x) - \lambda\varphi(x) \leq 0, \quad (t, x) \in J \times \mathbb{R}^d.$$

Conditions (i) and (ii) are standard regularity and ellipticity assumptions in parabolic PDE's. It is well known that assuming only (i) and (ii), problem (1.3) may admit several bounded solutions also in the autonomous case. Condition (iii) is mainly used to ensure *uniqueness* of the bounded classical solution u of (1.3) (i.e., uniqueness of a function $u \in C^{1,2}((s, +\infty) \times \mathbb{R}^d) \cap C_b([s, T] \times \mathbb{R}^d)$ for any $T > s$, that satisfies (1.3)).

In Section 2 we will be concerned with wellposedness of (1.3) in the space $C_b(\mathbb{R}^d)$. In the autonomous case the solutions to (1.3) are governed by a semigroup $\{T(t)\}$ which is the transition semigroup of the Markov process obtained in (1.1). In the non autonomous setting the semigroup is replaced by an evolution family $\{G(t, s)\}$. We will establish several properties of this family in Section 3. Note that, while regularity of $(G(t, s)\varphi)(x)$ with respect to (t, x) is a classical item in the theory of PDE's, regularity with respect to s is less standard. It is treated in the literature in the case of bounded coefficients because of its importance in several applications such as control theory. In our case, to get continuity with respect to s we have to sharpen Hypothesis 1.1(iii), assuming that $\mathcal{A}(t)\varphi$ is upperly bounded in $J \times \mathbb{R}^d$ for any bounded interval $J \subset I$.

In Section 4 we will study smoothing properties of $G(t, s)$, proving several estimates on the spatial derivatives of $G(t, s)\varphi$ for $\varphi \in C_b(\mathbb{R}^d)$. We will consider the following additional hypothesis:

- Hypothesis 1.2.**
- (i) The data q_{ij} and b_i ($i, j = 1, \dots, d$) and their first-order spatial derivatives belong to $C_{\text{loc}}^{\frac{\alpha}{2}, \alpha}(I \times \mathbb{R}^d)$;

(ii) there exists a continuous function $k : I \rightarrow \mathbb{R}$ such that

$$\langle \nabla_x b(t, x)\xi, \xi \rangle \leq k(t)|\xi|^2, \quad \xi \in \mathbb{R}^d, \quad (t, x) \in I \times \mathbb{R}^d;$$

(iii) there exists a continuous function $\rho : I \rightarrow [0, +\infty)$ such that, for every $i, j, k \in \{1, \dots, d\}$, we have

$$|D_k q_{ij}(t, x)| \leq \rho(t)\eta(t, x), \quad \xi \in \mathbb{R}^d, \quad (t, x) \in I \times \mathbb{R}^d.$$

Under this hypothesis we will prove uniform spatial gradient estimates for the function $G(t, s)f$ when $f \in C_b^k(\mathbb{R}^d)$, $k = 0, 1$, by means of the classical Bernstein method (see [2]). We will also prove more refined pointwise gradient estimates under either one of the following more restrictive conditions:

Hypothesis 1.3. (i) there exist a function $r : I \times \mathbb{R}^d \rightarrow \mathbb{R}$ and a constant $p_0 \in (1, +\infty)$ such that

$$\langle \nabla_x b(t, x)\xi, \xi \rangle \leq r(t, x)|\xi|^2, \quad \xi \in \mathbb{R}^d, \quad (t, x) \in I \times \mathbb{R}^d,$$

and

$$\sup_{(t,x) \in I \times \mathbb{R}^d} \left(r(t, x) + \frac{d^3(\rho(t))^2\eta(t, x)}{4 \min\{p_0 - 1, 1\}} \right) < +\infty;$$

(ii) Hypothesis 1.2(ii) holds true with the function k being replaced by a real constant k_0 . Moreover, there exists a positive constant ρ_0 such that, for every $i, j, k = 1, \dots, d$, we have

$$|D_k q_{ij}(t, x)| \leq \rho_0(\eta(t, x))^{\frac{1}{2}}, \quad (t, x) \in I \times \mathbb{R}^d.$$

Then, we get pointwise estimates,

$$(1.4) \quad |(\nabla G(t, s)f)(x)|^p \leq e^{\sigma_p(t-s)}(G(t, s)|\nabla f|^p)(x), \quad t \geq s, \quad x \in \mathbb{R}^d,$$

for every $p \geq p_0$ and some real constant σ_p . In the autonomous case (see [3]) these estimates are interesting not only for themselves, but also for the study of the behavior of the semigroup $\{T(t)\}$ in L^p -spaces with respect to invariant measures. An invariant measure corresponds to a stationary distribution of the Markov process with transition semigroup $\{T(t)\}$. In the analytical setting, an invariant measure for a Markov semigroup $\{T(t)\}$ is a Borel probability measure such that

$$\int_{\mathbb{R}^d} (T(t)f)(y) \mu(dy) = \int_{\mathbb{R}^d} f(y) \mu(dy), \quad t > 0, \quad f \in C_b(\mathbb{R}^d).$$

The interest in invariant measures is due to the following:

(i) the invariant measure arises naturally in the asymptotic behaviour of the semigroup. If μ is the (necessarily unique) invariant measure of $\{T(t)\}$, then

$$(T(t)f)(x) \rightarrow \int_{\mathbb{R}^d} f(y) \mu(dy) \quad \text{as } t \rightarrow +\infty,$$

for any $f \in C_b(\mathbb{R}^d)$ and any $x \in \mathbb{R}^d$;

(ii) the realizations of elliptic and parabolic operators in L^p -spaces with respect to invariant measures are dissipative.

In our nonautonomous case we cannot hope to find a single invariant measure. Instead, we look for systems of invariant measures (see e.g. [5, 7]) that is families of Borel probability measures $\{\mu_t : t \in I\}$ such that

$$\int_{\mathbb{R}^d} (G(t, s)f)(y) \mu_t(dy) = \int_{\mathbb{R}^d} f(y) \mu_s(dy), \quad t > s \in I, \quad f \in C_b(\mathbb{R}^d).$$

In Section 5, we will prove the existence of a system of invariant measures for our problem (1.3) replacing Hypothesis 1.1(iii) with the following stronger condition:

Hypothesis 1.4. there exist a nonnegative function $\varphi \in C^2(\mathbb{R}^d)$, diverging to $+\infty$ as $|x| \rightarrow +\infty$, and constants $a, c > 0$ and $t_0 \in I$ such that

$$(\mathcal{A}(t)\varphi)(x) \leq a - c\varphi(x), \quad t \geq t_0, \quad x \in \mathbb{R}^d.$$

In contrast to the autonomous case, systems of invariant measures are, in general, not unique. However, using a pointwise gradient estimate we will prove that uniqueness holds in the class of invariant measures $\{\mu_t : t \in I\}$ that admit finite moments of some order $p > 0$, which may blow up as $t \rightarrow +\infty$ with a certain exponential rate. By definition, $\{\mu_t : t \in I\}$ admits finite moments of order p if, for any $t \in I$,

$$\mu_t(p) = \int_{\mathbb{R}^d} |x|^p \mu_t(dx) < +\infty.$$

Still using a uniform gradient estimate, we show that, also in the non autonomous case, the asymptotic behaviour is determined by “the” system of invariant measures, in the sense that, for any $x \in \mathbb{R}^d$, $s \in I$, and $f \in C_b(\mathbb{R}^d)$,

$$\lim_{t \rightarrow +\infty} (G(t, s)f)(x) = \int_{\mathbb{R}^d} f(y) \mu_s(dy),$$

and the convergence is uniform in each compact set in \mathbb{R}^d .

Concerning point (ii), we note that, since we have to deal with a family of probability measures μ_t , we will also have a family of Lebesgue spaces $L^p(\mu_t)$ that are not mutually equivalent in general. This prevents us from extending the operators $G(t, s)$ to a single L^p -space, because $G(t, s)$ does not map $L^p(\mu_s)$ into itself in general, but it maps $L^p(\mu_s)$ into $L^p(\mu_t)$. However, it is possible to define an evolution semigroup associated with $G(t, s)$ on a single L^p -space of functions defined in $I \times \mathbb{R}^d$. This was already done in [6, 10, 11] in the special case of Ornstein-Uhlenbeck operators.

The evolution semigroup associated with an evolution family $\{G(t, s)\}$ is known to be a useful tool to determine several qualitative properties of the evolution family. See e.g. the book [4] and the references therein. In the case of time depending Ornstein-Uhlenbeck operators, the use of the evolution semigroup was essential to establish optimal regularity results for evolution equations and also to get precise asymptotic behavior estimates for $G(t, s)$, see [10, 11]. However, the general theory of evolution semigroups is well established only for evolution families acting on a fixed Banach space X , which is not our case. Therefore, the study of the asymptotic behavior of $G(t, s)\varphi$ for $\varphi \in L^p(\mu_s)$ through the evolution semigroup is deferred to a future paper. Here, we just describe the first properties of the evolution semigroup, in Section 6.

In the last section we consider a simple example and see how our conditions may be verified in this setting.

Notations. We denote, respectively, by $B_b(\mathbb{R}^d)$ and $C_b(\mathbb{R}^d)$ the set of all bounded and Borel measurable functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and its subset of all continuous functions. We endow both spaces with the sup-norm $\|\cdot\|_\infty$.

For any $k \in \mathbb{R}_+$ (possibly $k = +\infty$) we denote by $C_b^k(\mathbb{R}^d)$ the set of all functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ that are continuously differentiable in \mathbb{R}^d , up to $[k]$ -th-order, with bounded derivatives and such that the $[k]$ -th-order derivatives are $(k - [k])$ -Hölder continuous in \mathbb{R}^d . We norm $C_b^k(\mathbb{R}^d)$ by setting $\|f\|_{C_b^k(\mathbb{R}^d)} = \sum_{|\alpha| \leq [k]} \|D^\alpha f\|_\infty + \sum_{|\alpha|=[k]} [D^\alpha f]_{C_b^{k-[k]}(\mathbb{R}^d)}$. $C_c^k(\mathbb{R}^d)$ ($k \in \mathbb{N} \cup \{+\infty\}$) denotes the subset of $C_b^k(\mathbb{R}^d)$ of all compactly supported functions. $C_0(\mathbb{R}^d)$ denotes the set of all continuous functions vanishing at infinity.

If f is smooth enough we set $D_j f = \frac{\partial f}{\partial x_j}$,

$$|\nabla f(x)|^2 = \sum_{i=1}^d |D_i f(x)|^2, \quad |D^2 f(x)|^2 = \sum_{i,j=1}^d |D_{ij} f(x)|^2$$

and

$$\|\nabla f\|_\infty^2 = \sup_{x \in \mathbb{R}^d} |\nabla f(x)|^2, \quad \|D^2 f\|_\infty^2 = \sup_{x \in \mathbb{R}^d} |D^2 f(x)|^2.$$

Suppose that f depends on both time and spatial variables. If there is damage of confusion, we denote by $\nabla_x f$ and $D_x^2 f$ the gradient and the Hessian matrix of the function $f(t, \cdot)$. When f is a vector valued function, $\nabla_x f$ denotes the Jacobian matrix of $f(t, \cdot)$.

Let $D \subset \mathbb{R}^{d+1}$ be a domain or the closure of a domain. By $C_{\text{loc}}^{k+\alpha/2, 2k+\alpha}(D)$ ($k = 0, 1$, $\alpha \in (0, 1)$) we denote the set of all functions $f : D \rightarrow \mathbb{R}$ such that the time derivatives up to the k -th-order and the spatial derivatives up to the $2k$ -th-order are Hölder continuous with exponent α , with respect to the parabolic distance, in any compact set $D_0 \subset D$.

For any $r > 0$ we denote by $B_r \subset \mathbb{R}^d$ the open ball centered at 0 with radius r .

Given a measurable set E , we denote by $\mathbb{1}_E$ the characteristic function of E , i.e., $\mathbb{1}_E(x) = 1$ if $x \in E$, $\mathbb{1}_E(x) = 0$ otherwise.

Finally, we use the notation u_f for the (unique) bounded classical solution to problem (1.3).

2. SOLUTIONS IN $C_b(\mathbb{R}^d)$

In this section we want to solve our parabolic problem (1.3) with data $s \in I$ and $f \in C_b(\mathbb{R}^d)$. By a *solution* of (1.3) we mean a bounded classical solution, i.e. a function $u \in C_b([s, +\infty) \times \mathbb{R}^d) \cap C^{1,2}((s, +\infty) \times \mathbb{R}^d)$ such that (1.3) is satisfied. In the whole section we assume that Hypotheses 1.1 is fulfilled.

We already mentioned that Hypothesis 1.1(iii) ensures uniqueness of the solution to (1.3). In fact, it implies a maximum principle that we state as our first

Theorem 2.1. *Let $s \in I$ and $T > s$. If $u \in C_b([s, T] \times \mathbb{R}^d) \cap C^{1,2}((s, T] \times \mathbb{R}^d)$ satisfies*

$$\begin{cases} u_t(t, x) - \mathcal{A}(t)u(t, x) \leq 0, & (t, x) \in (s, T] \times \mathbb{R}^d, \\ u(s, x) \leq 0, & x \in \mathbb{R}^d, \end{cases}$$

then $u \leq 0$.

Proof. The proof can be obtained as the proof of [3, Theorem 4.1.5]. \square

Now, let us prove that problem (1.3) admits a unique bounded classical solution for any $f \in C_b(\mathbb{R}^d)$.

Theorem 2.2. *For every $s \in I$ and every $f \in C_b(\mathbb{R}^d)$, there exists a unique solution u of problem (1.3). Furthermore,*

$$(2.1) \quad \|u(t, \cdot)\|_\infty \leq \|f\|_\infty, \quad t \geq s.$$

Proof. Uniqueness follows from applying Theorem 2.1 to $u - v$ and to $v - u$, if u and v are two solutions. Estimate (2.1) follows by applying the same theorem to $\pm u - \|f\|_\infty$.

The existence part can be obtained in a classical way, solving Cauchy-Dirichlet problems in the balls B_n and then letting $n \rightarrow +\infty$. See e.g., [8, Proposition 2.2], [17, Theorem 4.2], [3, Theorems 2.2.1, 11.2.1]. Since there are some technicalities, for the reader's convenience we go into details. We split the proof in three steps.

Step 1. Here, we assume that f belongs to $C_c^{2+\alpha}(\mathbb{R}^d)$. Denote by n_0 the smallest integer such that $\text{supp}(f)$ is contained in the ball B_{n_0} . Further, for any $n \geq n_0$, we consider the Cauchy-Dirichlet problem

$$(2.2) \quad \begin{cases} u_t(t, x) = \mathcal{A}(t)u(t, x), & t \in (s, +\infty), \quad x \in B_n, \\ u(t, x) = 0, & t \in (s, +\infty), \quad x \in \partial B_n, \\ u(s, x) = f(x), & x \in B_n, \end{cases}$$

in the ball B_n . By classical results (see e.g., [9] or [15]) and Hypotheses 1.1(i)-(ii), for any $n \geq n_0$, the problem (2.2) admits a unique solution $u_n \in C_{\text{loc}}^{1+\frac{\alpha}{2}, 2+\alpha}([s, +\infty) \times \overline{B_n})$. Moreover, the classical Schauder estimates imply that, for any $m \in \mathbb{N}$, with $m > n_0$, there exists a constant $C = C(m)$ independent of n , such that

$$(2.3) \quad \|u_n\|_{C^{1+\frac{\alpha}{2}, 2+\alpha}(D_m)} \leq C\|f\|_{C^{2+\alpha}(\mathbb{R}^d)},$$

for any $n > m$, where $D_m = (s, m) \times B_m$. By the Arzelà-Ascoli theorem, there exists a subsequence (u_n^m) of u_n which converges in $C^{1,2}(\overline{D_m})$ to some function $u^m \in C^{1+\frac{\alpha}{2}, 2+\alpha}(D_m)$. Of course, u^m satisfies the differential equation $D_t u^m = \mathcal{A}(\cdot)u^m$ in D_m and it equals f on $\{s\} \times B_m$. Without loss of generality, we can assume that u_n^{m+1} is a subsequence of u_n^m . Note that, in this case, $u^{m+1}|_{D_m} \equiv u^m$. Hence, we can define a function u by putting $u|_{D_m} := u^m$. A standard procedure shows that u belongs to $C_{\text{loc}}^{1+\alpha/2, 2+\alpha}([s, +\infty) \times \mathbb{R}^d)$ and it is the classical solution to problem (1.3).

Note that the sequence u_n itself converges to u as n tends to $+\infty$, locally uniformly in $[s, +\infty) \times \mathbb{R}^d$. Indeed, the above arguments show that any convergent subsequence of (u_n) should converge to a classical solution of (1.3).

Step 2. Assume now that $f \in C_0(\mathbb{R}^d)$. Then, there exists a sequence $(f_n) \subset C_c^{2+\alpha}(\mathbb{R}^d)$ converging to f uniformly in \mathbb{R}^d as n tends to $+\infty$. Estimate (2.1) yields

$$\|u_{f_n} - u_{f_m}\|_{C_b([s, +\infty) \times \mathbb{R}^d)} \leq \|f_n - f_m\|_{C_b(\mathbb{R}^d)}, \quad m, n \in \mathbb{N}.$$

Therefore, there exists a bounded and continuous function u such that u_{f_n} converges to u , uniformly in $[s, +\infty) \times \mathbb{R}^d$. Moreover, applying the interior Schauder estimates to the sequence (u_{f_n}) , we deduce that u_{f_n} converges in $C_{\text{loc}}^{1,2}((s, +\infty) \times \mathbb{R}^d)$ to u . Hence, u is the bounded classical solution of problem (1.3).

Step 3. Now, fix $f \in C_b(\mathbb{R}^d)$ and consider a bounded sequence $(f_n) \in C_c^{2+\alpha}(\mathbb{R}^d)$ converging to f locally uniformly in \mathbb{R}^d as n tends to $+\infty$. The same arguments as in Step 2 show that, up to a subsequence, u_{f_n} converges, in $C_{\text{loc}}^{1,2}((s, +\infty) \times \mathbb{R}^d)$, to some function $u \in C_{\text{loc}}^{1+\alpha/2, 2+\alpha}((s, +\infty) \times \mathbb{R}^d)$, as n tends to $+\infty$. In particular, u solves the differential equation in (1.3). To prove that u is, actually, a classical solution of the problem (1.3), we fix a compact set $K \subset \mathbb{R}^d$ and a smooth and compactly supported function φ such that $0 \leq \varphi \leq 1$ and $\varphi \equiv 1$ in K . Further, we split $u_{f_n} = u_{\varphi f_n} + u_{(1-\varphi)f_n}$, for any $n \in \mathbb{R}$. Since the function φf is compactly supported in \mathbb{R}^d , it follows from Step 2 that $u_{\varphi f_n}$ converges to $u_{\varphi f}$ uniformly in $[s, +\infty) \times \mathbb{R}^d$.

Let us now consider the sequence $(u_{(1-\varphi)f_n})$. Fix $m \in \mathbb{N}$. We claim that

$$(2.4) \quad |(u_{(1-\varphi)f_m})(t, x)| \leq (1 - u_\varphi(t, x))M, \quad (t, x) \in (s, +\infty) \times \mathbb{R}^d,$$

where $M = \sup_{n \in \mathbb{N}} \|f_n\|_\infty$. Indeed, as a straightforward computation shows,

$$(1 - u_\varphi)M = u_{(1-\varphi)M}.$$

Therefore, the function $w := u_{(1-\varphi)f_m} - M(1 - u_\varphi)$ satisfies $w_t = \mathcal{A}w$ and, moreover,

$$w(s, \cdot) = (1 - \varphi)f_m - M(1 - \varphi) = (1 - \varphi)(f_m - M) \leq 0.$$

The maximum principle of Theorem 2.1 immediately implies that w is nonpositive in $[s, +\infty) \times \mathbb{R}^d$ or, equivalently, $u_{(1-\varphi)f_m} \leq (1-u_\varphi)M$. To prove the other inequality in (2.4), it suffices to observe that $(-u_{(1-\varphi)f_m}) = (u_{-(1-\varphi)f_m})$ and repeat the above arguments with f_m replaced by $-f_m$.

Now, since u_{f_n} converges pointwise to u , for any $(t, x) \in (s, +\infty) \times \mathbb{R}^d$ we have

$$|u(t, x) - f(x)| = \lim_{n \rightarrow +\infty} |u_{f_n}(t, x) - f(x)|, \quad (t, x) \in (s, +\infty) \times \mathbb{R}^d,$$

and, for each $n \in \mathbb{N}$, we have

$$\begin{aligned} |u_{f_n}(t, x) - f(x)| &\leq |u_{\varphi f_n}(t, x) - f(x)| + |u_{(1-\varphi)f_n}(t, x)| \\ (2.5) \quad &\leq |u_{\varphi f_n}(t, x) - f(x)| + (1-u_\varphi(t, x))M. \end{aligned}$$

The right-hand side of (2.5) converges to 0 uniformly in K as t tends to s^+ . Hence, u can be continuously extended up to $t = s$ setting $u(s, \cdot) = f$. This completes the proof. \square

Remark 2.3. Let us observe that the choice of approximating problem (1.3) by Cauchy-Dirichlet problems in the ball B_n is not essential. Indeed, repeating step by step the proof of Theorem 2.2, we can see that problem (1.3) can be approximated also by the Cauchy-Neumann problems

$$(2.6) \quad \begin{cases} u_t(t, x) = \mathcal{A}(t)u(t, x), & (t, x) \in (s, +\infty) \times B_n, \\ \frac{\partial}{\partial \nu}u(t, x) = 0, & (t, x) \in (s, +\infty) \times \partial B_n, \\ u(s, x) = f(x), & x \in B_n. \end{cases}$$

We will use this approach in Section 4 to prove estimates for the space derivatives of $G(t, s)f$.

Now we define the *evolution family* associated with our problem (1.3). Let

$$\Lambda := \{(t, s) \in I \times I : t \geq s\}.$$

We put $G(t, t) := id_{C_b(\mathbb{R}^d)}$ and for $t > s$ we define the operator $G(t, s)$ by setting

$$(G(t, s)f)(x) := u_f(t, x), \quad x \in \mathbb{R}^d,$$

where u_f is the unique solution of problem (1.3). We call the family $\{G(t, s) : (t, s) \in \Lambda\}$ the *evolution family associated with the problem (1.3)*.

It is immediate from Theorem 2.1 that, for $(t, s) \in \Lambda$, the operator $G(t, s)$ is a positive contraction on $C_b(\mathbb{R}^d)$. From the uniqueness assertion in Theorem 2.2, the *law of evolution*

$$(2.7) \quad G(t, s)G(s, r) = G(t, r),$$

for $r \leq s \leq t$ easily follows.

The connection with Markov processes suggests that every operator $G(t, s)$ should be associated with a transition kernel. Recall that a transition kernel p is a mapping from $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d)$ such that $p(x, \cdot)$ is a sub probability measure for fixed x and such that $p(\cdot, A)$ is measurable for fixed $A \in \mathcal{B}(\mathbb{R}^d)$. The following proposition states that this is indeed the case; in fact, the transition kernels $p_{t,s}$ form the non autonomous equivalent of a conservative, stochastically continuous transition function, cf. [7, Sections 2.1 and 2.8].

Proposition 2.4. *For every $(t, s) \in \Lambda$ and every $x \in \mathbb{R}^d$ there exists a unique probability measure $p_{t,s}(x, \cdot)$ such that*

$$(2.8) \quad (G(t, s)f)(x) = \int_{\mathbb{R}^d} f(y) p_{t,s}(x, dy),$$

for each $C_b(\mathbb{R}^d)$. Furthermore, the following properties hold:

- (i) for every $t \in I$, $p_{t,t}(x, \cdot)$ is the Dirac measure concentrated at x ;
- (ii) for $t > s$ the measure $p_{t,s}(x, \cdot)$ is equivalent to the Lebesgue measure, i.e. they have the same sets of zero measure;
- (iii) for fixed $A \in B_b(\mathbb{R}^d)$ and $(t, s) \in \Lambda$ the map $x \mapsto p_{t,s}(x, A)$ is Borel measurable;
- (iv) for every open set $U \subset \mathbb{R}^d$ containing x we have

$$\lim_{t \rightarrow s^+} p_{t,s}(x, U) = 1;$$

- (v) for every $t \geq s \geq r$, $x \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$ we have

$$p_{t,r}(x, A) = \int_{\mathbb{R}^d} p_{s,r}(y, A) p_{t,s}(x, dy).$$

Proof. Let us define

$$p_{t,s}(x, A) = \int_A g(t, s, x, y) dy,$$

for any (Lebesgue) measurable set $A \subset \mathbb{R}^d$ and any $t > s$, where g is the Green function of problem (1.3) which can be obtained as the pointwise limit of the increasing (with respect to n) sequence of Green functions g_n associated with problem (2.2). For the existence of these latter kernels, see e.g., [9, Theorem 3.16]. The function g is measurable in its entries and it is positive since the g_n 's are.

Notice that, since $G(t, s)\mathbf{1} \equiv \mathbf{1}$ by uniqueness, we have

$$p_{t,s}(x, \mathbb{R}^d) = \int_{\mathbb{R}^d} \mathbf{1}(y) p_{t,s}(x, dy) = (G(t, s)\mathbf{1})(x) = 1,$$

i.e. $p_{t,s}(x, \cdot)$ is a probability measure. Now formula (2.8) and properties (i) and (ii) immediately follow.

To prove (iii), let $A \in \mathcal{B}(\mathbb{R}^d)$. Then, there exists a bounded sequence $(f_n) \subset C_b(\mathbb{R}^d)$ converging almost everywhere to $\mathbf{1}_A$. Hence, by the dominated convergence theorem,

$$\begin{aligned} p_{t,s}(x, A) &= \int_{\mathbb{R}^d} \mathbf{1}_A(y) p_{t,s}(x, dy) \\ &= \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^d} f_n(y) p_{t,s}(x, dy) \\ &= \lim_{n \rightarrow +\infty} (G(t, s)f_n)(x), \end{aligned}$$

for any $(s, t) \in \Lambda$ and any $x \in \mathbb{R}^d$, and this implies that the function $(t, s, x) \mapsto p_{t,s}(x, A)$ is measurable.

Property (iv) follows from the continuity of the map $G(\cdot, s)f$ on $\{s\} \times \mathbb{R}^d$, by virtue of [7, Lemma 2.2].

Finally, (v) is an immediate consequence of (2.7). \square

Corollary 2.5. *The evolution family $\{G(t, s)\}$ is irreducible, i.e., $(G(t, s)\mathbf{1}_U)(x) > 0$ for any open set $U \subset \mathbb{R}^d$, any $(s, t) \in \Lambda$ and any $x \in \mathbb{R}^d$. More generally, $(G(t, s)\mathbf{1}_A)(x) > 0$ for each Borel set $A \subset \mathbb{R}^d$ with positive Lebesgue measure, and for any $(s, t) \in \Lambda$ and any $x \in \mathbb{R}^d$.*

Proof. The statement follows from the inequality $g(t, s, x, y) > 0$ $((s, t) \in \Lambda, x \in \mathbb{R}^d)$ which holds since the Green function g is the pointwise limit of the increasing sequence of Green functions g_n associated with problem (2.2). \square

Remark 2.6. The representation (2.8) implies that we can extend our evolution family $G(t, s)$ to the space $B_b(\mathbb{R}^d)$ of bounded, measurable functions. More generally,

in the sequel we set

$$(G(t, s)f)(x) := \int_{\mathbb{R}^d} f(y)p_{t,s}(x, dy), \quad x \in \mathbb{R}^d,$$

also for unbounded functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$, provided that $f \in L^1(\mathbb{R}^d, p_{t,s}(x, \cdot))$.

Formula (2.8) also implies that the adjoints $G^*(t, s)$ leave the space of signed measures invariant.

3. CONTINUITY PROPERTIES OF THE EVOLUTION FAMILY $\{G(t, s)\}$

In this section we prove some useful continuity properties of the function $G(t, s)f$ when $f \in C_b(\mathbb{R}^d)$. To begin with, let us prove the following proposition.

Proposition 3.1. *Let $(f_n) \subset C_b(\mathbb{R}^d)$ be a sequence such that $\|f_n\|_\infty \leq M$ for each $n \in \mathbb{N}$. Then:*

- (i) *if f_n converges pointwise to f , then $G(\cdot, s)f_n$ converges to $G(\cdot, s)f$ uniformly in $[s + \varepsilon, T] \times K$ for each $T > s + \varepsilon > s$ and each compact set $K \subset \mathbb{R}^d$;*
- (ii) *if f_n converges uniformly to f in each compact set $K \subset \mathbb{R}^d$, then $G(\cdot, s)f_n$ converges to $G(\cdot, s)f$ uniformly in $[s, T] \times K$, for each $T > s$ and each compact set $K \subset \mathbb{R}^d$.*

Proof. (i). By the representation formula (2.8) and the dominated convergence theorem, we obtain that $G(\cdot, s)f_n$ converges to $G(\cdot, s)f$ pointwise in $[s, +\infty) \times \mathbb{R}^d$, as n tends to $+\infty$. However, the classical interior Schauder estimates yield

$$\|G(\cdot, s)f_n\|_{C^{1+\alpha/2, 2+\alpha}([s+\varepsilon, T] \times K)} \leq C\|f_n\|_\infty \leq CM, \quad n \in \mathbb{N},$$

for any $\varepsilon > 0$, any compact set $K \subset \mathbb{R}^d$ and some positive constant $C = C(\varepsilon, K)$. The Arzelà-Ascoli theorem implies that $G(\cdot, s)f_n$ converges to $G(\cdot, s)f$ in $C^{1,2}([s + \varepsilon, T] \times K)$.

(ii). The proof follows adapting the arguments used in the proof of Theorem 2.2. Let K be a compact set and let $\varphi \in C_c^{2+\alpha}(\mathbb{R}^d)$ be such that $\varphi \equiv 1$ in K . Split $u_{f_n} = u_{\varphi f_n} + u_{(1-\varphi)f_n}$. By Step 2 in the proof of Theorem 2.2, $u_{\varphi f_n}$ converges to $u_{\varphi f}$ uniformly in $[s, +\infty) \times \mathbb{R}^d$.

To complete the proof, it suffices to show that $u_{(1-\varphi)f_n}$ converges to $u_{(1-\varphi)f}$ uniformly in $[s, T] \times K$ as n tends to $+\infty$ for any $T > s$. The arguments in Step 2 of the proof of Theorem 2.2 show that, up to a subsequence, $u_{(1-\varphi)f_n}$ converges in $C^{1,2}([s + \varepsilon, T] \times B_r)$, for any $\varepsilon > 0$ and any $r > 0$, to a function $v \in C_{loc}^{1+\alpha/2, 2+\alpha}((s, +\infty) \times \mathbb{R}^d)$. Moreover, letting m go to $+\infty$ in (2.4) gives

$$(3.1) \quad |v(t, x)| \leq (1 - u_\varphi(t, x))M, \quad (t, x) \in (s, +\infty) \times \mathbb{R}^d.$$

Since u_φ is continuous at $\{s\} \times \mathbb{R}^d$ and $\varphi \equiv 1$ in K , from (3.1) we deduce that $v(t, x)$ converges to 0 as $t \rightarrow s^+$, uniformly with respect to $x \in K$.

Let us now fix $\varepsilon > 0$ and let δ be sufficiently small such that $|u_{(1-\varphi)f_n}| + |v| \leq \varepsilon/2$ in $[s, s + \delta] \times K$ for any $n \in \mathbb{N}$. Moreover, we fix n large enough such that $|u_{(1-\varphi)f_n} - v| \leq \varepsilon/2$ in $[s + \delta, T] \times K$. For such n and δ we get

$$\begin{aligned} \|u_{(1-\varphi)f_n} - v\|_{C([s, T] \times K)} &\leq \|u_{(1-\varphi)f_n}\|_{C([s, s + \delta] \times K)} + \|v\|_{C([s, s + \delta] \times K)} \\ &\quad + \|u_{(1-\varphi)f_n} - v\|_{C[s + \delta, T] \times K)} \\ &\leq \varepsilon. \end{aligned}$$

Summing up, the sequence $u_{f_n} = u_{\varphi f_n} + u_{(1-\varphi)f_n}$ converges as n tends to $+\infty$, to the function $u = u_{\varphi f} + v$ which belongs to $C_{loc}^{1+\alpha/2, 2+\alpha}((s, +\infty) \times \mathbb{R}^d)$, and the convergence is uniform in $[s, T] \times K$. Since K is arbitrary, u_{f_n} converges locally uniformly to u in $[s, +\infty) \times \mathbb{R}^d$, so that u is continuous up to $t = s$ where it equals f . Moreover, since u_{f_n} converges to u in $C^{1,2}([s + \varepsilon, T] \times B_R)$ for any $\varepsilon \in (0, T - s)$

and any $R > 0$, then $D_t u - \mathcal{A}u = 0$ for $t > s$. Thus, u is a bounded classical solution of (1.3) and, by Theorem 2.1, $u = u_f$. This completes the proof. \square

3.1. Continuity of the function $G(t, s)f$ with respect to the variable s . Since evolution families depend on two parameters t and s , it is natural to investigate also the smoothness of the function $G(t, \cdot)f$. In the following lemma we prove a very useful generalization of the well known formula that holds in the case of bounded coefficients. This lemma will play a fundamental role to prove existence of evolution systems of invariant measures in Section 5.

Lemma 3.2. *Let $f \in C_b^2(\mathbb{R}^d)$ be constant outside a compact set K . Then, for any $x \in \mathbb{R}^d$, and any $s_0 < s_1 \leq t$, the function $r \mapsto (G(t, r)\mathcal{A}(r)f)(x)$ is integrable in (s_0, s_1) and we have*

$$(3.2) \quad (G(t, s_1)f)(x) - (G(t, s_0)f)(x) = - \int_{s_0}^{s_1} (G(t, r)\mathcal{A}(r)f)(x) dr.$$

In particular, the function $(G(t, \cdot)f)(x)$ is continuously differentiable in $I_t := I \cap (-\infty, t]$ and

$$(3.3) \quad \frac{d}{ds}(G(t, s)f)(x) = -(G(t, s)\mathcal{A}(s)f)(x), \quad s \in I_t.$$

Finally for any $g \in C_0(\mathbb{R}^d)$ the function $G(t, \cdot)g$ is continuous in I_t with values in $C_b(\mathbb{R}^d)$.

Proof. By assumption, we can write $f = g + c \cdot \mathbf{1}$, for some $g \in C_c^2(\mathbb{R}^d)$ and some $c \in \mathbb{R}$. However, $G(t, s)\mathbf{1} \equiv \mathbf{1}$, whence the assertion is trivially satisfied by any constant function. Thus, it remains to prove it when $f \in C_c^2(\mathbb{R}^d)$. Choose n_0 such that $\text{supp}(f) \subset B_{n_0}$, and denote by $\{G_n(t, s)\}$ the evolution family associated with problem (2.2) for $n \geq n_0$ (cf. [1, Theorem 6.3]). By [1, Theorem 2.3(ix)], we can write (3.3) with G being replaced by G_n . Integrating such an equality with respect to s and recalling that, by Step 1 in the proof of Theorem 2.2, for any $(t, r) \in \Lambda$, $G_n(t, r)f$ converges to $G(t, r)f$ pointwise in \mathbb{R}^d as n tends to $+\infty$, we obtain

$$(3.4) \quad \begin{aligned} (G(t, s_1)f)(x) - (G(t, s_0)f)(x) &= \lim_{n \rightarrow +\infty} (G_n(t, s_1)f)(x) - (G_n(t, s_0)f)(x) \\ &= - \lim_{n \rightarrow +\infty} \int_{s_0}^{s_1} (G_n(t, r)\mathcal{A}(r)f)(x) dr \\ &= - \int_{s_0}^{s_1} (G(t, r)\mathcal{A}(r)f)(x) dr, \end{aligned}$$

where the last equality follows by dominated convergence.

Now, observe that (3.4) implies that the function $G(t, \cdot)f$ is continuous in I_t , with values in $C_b(\mathbb{R}^d)$, for any $f \in C_c^2(\mathbb{R}^d)$. Since $C_c^2(\mathbb{R}^d)$ is dense in $C_0(\mathbb{R}^d)$, $G(t, \cdot)g$ is continuous in I_t for any $f \in C_0(\mathbb{R}^d)$.

To prove that the function $G(t, \cdot)f$ is differentiable, it is enough to show that the function $G(t, \cdot)\mathcal{A}(\cdot)f$ is continuous in I_t . Indeed, for any $r, r_0 \in I_t$,

$$\begin{aligned} &\|G(t, r)\mathcal{A}(r)f - G(t, r_0)\mathcal{A}(r_0)f\|_\infty \\ &= \|G(t, r)(\mathcal{A}(r)f - \mathcal{A}(r_0)f)\|_\infty + \|(G(t, r) - G(t, r_0))\mathcal{A}(r_0)f\|_\infty \\ &\leq \|\mathcal{A}(r)f - \mathcal{A}(r_0)f\|_\infty + \|(G(t, r) - G(t, r_0))\mathcal{A}(r_0)f\|_\infty, \end{aligned}$$

and the last side of the previous chain of inequalities goes to 0 as $r \rightarrow r_0$, since $\mathcal{A}(r_0)f \in C_c(\mathbb{R}^d)$. Now, (3.4) implies that the function $G(t, \cdot)f$ is differentiable in I_t , and (3.3) follows. This completes the proof. \square

To prove that $(t, s, x) \mapsto (G(t, s)f)(x)$ is continuous in $\Lambda \times \mathbb{R}^d$ for any function $f \in C_b(\mathbb{R}^d)$, we need an intermediate assumption between Hypothesis 1.1(iii) and

Hypothesis 1.4. More precisely, in the rest of this section we assume that the following hypothesis is satisfied.

Hypothesis 3.3. For every bounded interval $J \subset I$ there exist a function $\varphi = \varphi_J \in C^2(\mathbb{R}^d)$ diverging to $+\infty$ as $|x|$ tends to $+\infty$, and a positive constant M_J such that

$$(\mathcal{A}(t)\varphi)(x) \leq M_J, \quad t \in J, \quad x \in \mathbb{R}^d.$$

Hypothesis 3.3 allows to define $G(t, s)$ on a larger class than $B_b(\mathbb{R}^d)$. Namely, we show that the right hand side of (2.8) makes sense for $f = \varphi$, where φ is any of the functions in Hypothesis 3.3.

Let us begin with the following fundamental lemma. If $J \subset I$ is any interval, we set

$$\Lambda_J := \{(t, s) \in J \times J : s \leq t\}.$$

Lemma 3.4. Assume that Hypotheses 1.1(i)(ii) and 3.3 are satisfied. Fix a bounded interval $J \subset I$ and let $\varphi = \varphi_J$ be as in Hypothesis 3.3. Then, the function $(t, s, x) \mapsto (G(t, s)\varphi)(x)$ is well defined and bounded in $\Lambda_J \times B_\varrho$, for every $\varrho > 0$.

Proof. We may assume (possibly adding a constant) that $\varphi(x) \geq 0$ for each $x \in \mathbb{R}^d$.

For every $n \in \mathbb{N}$ choose a function $\psi_n \in C^\infty([0, +\infty))$ such that

- (i) $\psi_n(t) = t$ for $t \in [0, n]$,
- (ii) $\psi_n(t) \equiv \text{const.}$ for $t \geq n+1$,
- (iii) $0 \leq \psi'_n \leq 1$ and $\psi''_n \leq 0$.

Then, the function $\varphi_n := \psi_n \circ \varphi$ belongs to $C_b^2(\mathbb{R}^d)$ and it is constant outside a compact set. By Lemma 3.2, we have

$$\begin{aligned} \varphi_n(x) &\geq \varphi_n(x) - (G(t, s)\varphi_n)(x) \\ &= - \int_s^t \int_{\mathbb{R}^d} (\mathcal{A}(r)\varphi_n)(y) p_{t,r}(x, dy) dr \\ &= - \int_s^t \int_{\mathbb{R}^d} \{\psi'_n(\varphi)(\mathcal{A}(r)\varphi)(y) + \psi''_n(\varphi) \langle Q(r, y) \nabla \varphi(y), \nabla \varphi(y) \rangle\} p_{t,r}(x, dy) dr \\ (3.5) \quad &\geq - \int_s^t \int_{\mathbb{R}^d} \psi'_n(\varphi)(\mathcal{A}(r)\varphi)(y) p_{t,r}(x, dy) dr, \end{aligned}$$

for any $s, t \in \Lambda$ and any $x \in \mathbb{R}^d$. We claim that for each $s, t \in J$, letting $n \rightarrow +\infty$ in (3.5) we obtain

$$\varphi(x) \geq - \int_s^t \int_{\mathbb{R}^d} (\mathcal{A}(r)\varphi)(y) p_{t,r}(x, dy) dr = - \int_s^t (G(t, r)\mathcal{A}(r)\varphi)(x) dr,$$

so that, in particular, the above integral is finite.

It is clear that $\lim_{n \rightarrow +\infty} \varphi_n(x) = \varphi(x)$ for each $x \in \mathbb{R}^d$. Concerning the integral in the right-hand side of (3.5), we split it into the sum

$$\begin{aligned} &\int_s^t \int_{\mathbb{R}^d} \psi'_n(\varphi)(\mathcal{A}(r)\varphi)(y) p_{t,r}(x, dy) dr \\ &= - \int_s^t \int_{\mathbb{R}^d} \psi'_n(\varphi) \{M_J - (\mathcal{A}(r)\varphi)(y)\} p_{t,r}(x, dy) dr \\ (3.6) \quad &\quad + M_J \int_s^t \int_{\mathbb{R}^d} \psi'_n(\varphi) p_{t,r}(x, dy) dr. \end{aligned}$$

Since $\psi'_n(\varphi)(y)$ is increasing in n and converges to 1 for each y , both integrals in the right-hand side of (3.6) converge by the monotone convergence theorem. The claim follows.

Letting $n \rightarrow +\infty$ in (3.5) yields

$$(3.7) \quad (G(t, s)\varphi)(x) \leq \varphi(x) + \int_s^t (G(t, r)\mathcal{A}(r)\varphi)(x) dr,$$

and since $(\mathcal{A}(r)\varphi)(y) \leq M_J$ for each $y \in \mathbb{R}^d$ and $r \in J$,

$$(3.8) \quad \int_s^t G(t, r)(\mathcal{A}(r)\varphi)(y) dr \leq M_J(t - s).$$

Estimates (3.7) and (3.8) imply that

$$(G(t, s)\varphi)(x) \leq \varphi(x) + M_J(t - s),$$

for any $s, t \in J$, with $s \leq t$ and any $x \in \mathbb{R}^d$. It follows that

$$(3.9) \quad M_{J,\varrho} := \sup_{\substack{(t,s,x) \in J \times J \times \mathbb{R}^d \\ s \leq t, |x| \leq \varrho}} (G(t, s)\varphi)(x) < +\infty.$$

This completes the proof. \square

Having $(G(t, s)\varphi)(x)$ bounded for $(t, s) \in \Lambda_J$, we may prove in the standard way that for each $r > 0$ the family of measures $\{p_{t,s}(x, dy) : (t, s, x) \in \Lambda_J \times \overline{B_r}\}$ is tight. We recall that a family of (probability) measures $\{\mu_\alpha : \alpha \in \mathcal{F}\}$ is tight, if for any $\varepsilon > 0$ there exists $\varrho > 0$ such that $\mu_\alpha(\mathbb{R}^d \setminus B(\varrho)) \leq \varepsilon$ for any $\alpha \in \mathcal{F}$.

Lemma 3.5. *Under the assumptions of Lemma 3.4, for each bounded interval $J \subset I$ and for each $r > 0$ the family of measures $\{p_{t,s}(x, dy) : (t, s, x) \in \Lambda_J \times \overline{B_r}\}$ is tight.*

Proof. Fix $\varepsilon > 0$ and consider the function $\varphi = \varphi_J$ in Hypothesis 3.3. As in the proof of Lemma 3.4, we assume that φ is nonnegative. Since φ blows up as $|x| \rightarrow +\infty$, there exists $\varrho > 0$ such that

$$\varphi(x) \geq \frac{M_{J,r}}{\varepsilon} (\mathbf{1}_{\mathbb{R}^d \setminus B(\varrho)})(x), \quad x \in \mathbb{R}^d,$$

where $M_{J,r}$ is given by (3.9). Then, for $(t, s) \in \Lambda_J$, we have

$$\begin{aligned} p_{t,s}(x, \mathbb{R}^d \setminus B(\varrho)) &= \int_{\mathbb{R}^d} \mathbf{1}_{\mathbb{R}^d \setminus B(\varrho)}(y) p_{t,s}(x, dy) \\ &\leq \frac{\varepsilon}{M_{J,r}} \int_{\mathbb{R}^d} \varphi(y) p_{t,s}(x, dy) \\ &= \frac{\varepsilon}{M_{J,r}} (G(t, s)\varphi)(x) \leq \varepsilon, \end{aligned}$$

so that

$$(3.10) \quad \sup_{(t,s,x) \in \Lambda_J \times \overline{B_r}} p_{t,s}(x, \mathbb{R}^d \setminus B(\varrho)) \leq \varepsilon,$$

and the statement follows. \square

As usual, tightness yields some convergence result.

Proposition 3.6. *Assume that Hypotheses 1.1(i)-(ii) and 3.3 are satisfied. Further, let $\{f_n\}$ be a bounded sequence in $C_b(\mathbb{R}^d)$, such that $\|f_n\|_\infty \leq M$ for each $n \in \mathbb{N}$ and f_n converges to $f \in C_b(\mathbb{R}^d)$ locally uniformly in \mathbb{R}^d . Then, the function $G(\cdot, \cdot)f_n$ converges to $G(\cdot, \cdot)f$ locally uniformly in $\Lambda \times \mathbb{R}^d$.*

Proof. Fix any bounded interval $J \subset I$ and any $\varepsilon, r > 0$. Let ϱ be such that (3.10) holds, and for $(t, s, x) \in \Lambda_J \times \overline{B_\rho}$ split $G(t, s)f_n - G(t, s)f$ as

$$\begin{aligned} (G(t, s)f_n)(x) - (G(t, s)f)(x) &= \int_{\mathbb{R}^d} (f_n(y) - f(y))p_{t,s}(x, dy) \\ &= \int_{B(\varrho)} (f_n(y) - f(y))p_{t,s}(x, dy) \\ &\quad + \int_{\mathbb{R}^d \setminus B(\varrho)} (f_n(y) - f(y))p_{t,s}(x, dy), \end{aligned}$$

so that

$$\begin{aligned} |(G(t, s)f_n)(x) - (G(t, s)f)(x)| &\leq \sup_{y \in B(\varrho)} |f_n(y) - f(y)| \int_{\mathbb{R}^d} p_{t,s}(x, dy) \\ &\quad + \left(\sup_{n \in \mathbb{N}} \|f_n\|_\infty + \|f\|_\infty \right) \int_{\mathbb{R}^d \setminus B(\varrho)} p_{t,s}(x, dy) \\ &\leq \sup_{y \in B(\varrho)} |f_n(y) - f(y)| \\ &\quad + 2M\varepsilon. \end{aligned}$$

Fix $n_0 \in \mathbb{N}$ such that

$$\sup_{y \in B(\varrho)} |f_n(y) - f(y)| \leq \varepsilon, \quad n \geq n_0.$$

For $n \geq n_0$ we get

$$\begin{aligned} &\sup_{(t,s,x) \in \Lambda_J \times \overline{B_r}} |(G(t, s)f_n)(x) - (G(t, s)f)(x)| \\ &= \varepsilon(1 + 2M). \end{aligned}$$

Thus, $G(\cdot, \cdot)f_n$ converges to $G(\cdot, \cdot)f$ uniformly in $\Lambda_J \times \overline{B_r}$. \square

Now we are ready to prove that $(t, s, x) \mapsto (G(t, s)f)(x)$ is continuous, for each $f \in C_b(\mathbb{R}^d)$.

Theorem 3.7. *Under the assumptions of Proposition 3.6, the function $(t, s, x) \mapsto (G(t, s)f)(x)$ is continuous in $\Lambda \times \mathbb{R}^d$, for any $f \in C_b(\mathbb{R}^d)$.*

Proof. Fix $f \in C_b(\mathbb{R}^d)$ and let $\{f_n\} \in C_c^\infty(\mathbb{R}^d)$ be a sequence of smooth functions converging to f locally uniformly in \mathbb{R}^d and such that

$$\sup_{n \in \mathbb{N}} \|f_n\|_\infty < +\infty.$$

By Proposition 3.6, the sequence of functions $(t, s, x) \mapsto (G(t, s)f_n)(x)$ converges to $(t, s, x) \mapsto (G(t, s)f)(x)$ locally uniformly. Therefore, it suffices to show that $(t, s, x) \mapsto (G(t, s)g)(x)$ is continuous in $\Lambda \times \mathbb{R}^d$ whenever $g \in C_c^\infty(\mathbb{R}^d)$. For this purpose, we observe that the classical interior Schauder estimates as in [9, Theorem 3.5] imply a slightly more general estimate than (2.3), i.e.,

$$(3.11) \quad \sup_{s \in [a, b]} \|G_n(\cdot, s)g\|_{C^{1+\alpha/2, 2+\alpha}([s, s+m] \times B(m))} \leq C\|g\|_{C_b^{2+\alpha}(\mathbb{R}^d)},$$

for any $a, b \in I$, $a < b$, and some positive constant C , independent of $n > m$.

Since the sequence of functions $(t, x) \mapsto (G_n(t, s)g)(x)$ converges to $(t, x) \mapsto (G(t, s)g)(x)$ in $C^{1+\alpha/2, 2+\alpha}([s, s+m] \times B(m))$ for any $s \in [a, b]$, it follows that $(t, x) \mapsto (G(t, s)g)(x) \in C^{1+\alpha/2, 2+\alpha}([s, s+m] \times B(m))$ for any $s \in [a, b]$ and its $C^{1+\alpha/2, 2+\alpha}$ -norm is bounded by $C\|g\|_{C_b^{2+\alpha}(\mathbb{R}^d)}$, with the constant C of formula (3.11).

Fix (t_0, s_0, x_0) , $(t, s, x) \in \Lambda \times \mathbb{R}^d$, with $s_0, s \in [a, b]$. Suppose that $s_0 \leq s$. Then, $(t, s_0) \in \Lambda$, and

$$(3.12) \quad \begin{aligned} |(G(t, s)g)(x) - (G(t_0, s_0)g)(x_0)| &\leq |(G(t, s)g)(x) - (G(t, s_0)g)(x)| \\ &\quad + |(G(t, s_0)g)(x) - (G(t_0, s_0)g)(x_0)|. \end{aligned}$$

By (3.2) there exists a positive constant $C = C(a, b)$ such that

$$(3.13) \quad |(G(t, s)g)(x) - (G(t, s_0)g)(x)| \leq C|s - s_0|.$$

Combining (3.12) and (3.13) yields

$$\lim_{\substack{(t, s, x) \rightarrow (t_0, s_0, x_0) \\ s \geq s_0}} (G(t, s)g)(x) = (G(t_0, s_0)g)(x_0).$$

Let now assume that $s < s_0$ and split

$$(3.14) \quad \begin{aligned} |(G(t, s)g)(x) - (G(t_0, s_0)g)(x_0)| &\leq |(G(t, s)g)(x) - (G(t_0, s)g)(x_0)| \\ &\quad + |(G(t_0, s)g)(x_0) - (G(t_0, s_0)g)(x_0)|. \end{aligned}$$

Since $(t, x) \mapsto (G(\cdot, s)g)(x)$ is continuous in $[s, +\infty) \times \mathbb{R}^d$, locally uniformly with respect to s , from (3.13) and (3.14) we also deduce that

$$\lim_{\substack{(t, s, x) \rightarrow (t_0, s_0, x_0) \\ s < s_0}} (G(t, s)g)(x) = (G(t_0, s_0)g)(x_0).$$

This completes the proof. \square

4. GRADIENT ESTIMATES

In this section we prove both uniform and pointwise gradient estimates. Besides being interesting in their own right, we will need them in the next section to prove uniqueness of systems of invariant measures in a suitable class and convergence results.

Throughout the section we assume that Hypotheses 1.1 and 1.2 hold. Therefore, the bounded classical solution of problem (1.2) is such that its first-order spatial derivatives belong to $C_{\text{loc}}^{1+\frac{\alpha}{2}, 2+\alpha}((s, +\infty) \times \mathbb{R}^d)$ (see e.g., [9, Theorem 3.10] and [16]). We will use this fact in the sequel to apply a variant of the Bernstein method to get our gradient estimates.

First, we prove uniform gradient estimates.

Theorem 4.1. *Let $s \in I$ and $T > s$. Then, there exist positive constants C_1, C_2 , depending on s and T , such that:*

(i) *for every $f \in C_b^1(\mathbb{R}^d)$ we have*

$$\|\nabla G(t, s)f\|_\infty \leq C_1 \|f\|_{C_b^1(\mathbb{R}^d)}, \quad s < t \leq T;$$

(ii) *for every $f \in C_b(\mathbb{R}^d)$ we have*

$$\|\nabla G(t, s)f\|_\infty \leq \frac{C_2}{\sqrt{t-s}} \|f\|_\infty, \quad s < t \leq T.$$

Proof. It suffices to prove the statement for $f \in C_c^{2+\alpha}(\mathbb{R}^d)$, since we may approximate an arbitrary f by a sequence $(f_n) \subset C_c^{2+\alpha}(\mathbb{R}^d)$, bounded with respect to the sup-norm and converging to f locally uniformly in \mathbb{R}^d , and Step 3 of Theorem 2.2 shows that $\nabla G(\cdot, s)f_n$ converge to $\nabla G(\cdot, s)f$ pointwise in $(s, T] \times \mathbb{R}^d$.

Let k, ρ be the functions in Hypothesis 1.2. Set

$$k_0 := \sup_{t \in [s, T]} k(t), \quad \rho_0 := \sup_{t \in [s, T]} \rho(t).$$

(i). Let u_n be the unique solution of the Cauchy-Neumann problem (2.6), where n is so large that the support of f is contained in B_n . By Remark 2.3, u_n converges

to $u(t, x) := (G(t, s)f)(x)$ in $C^{1,2}([s, T] \times K)$ as $n \rightarrow +\infty$, for any compact set $K \subset \mathbb{R}^d$.

Define

$$z_n(t, x) = u_n(t, x)^2 + a|\nabla_x u_n(t, x)|^2, \quad (t, x) \in [s, T] \times B_n.$$

Then, z_n belongs to $C^{1,2}((s, +\infty) \times B_n) \cap C_b([s, T] \times \overline{B_n})$, for any $s < T$. Since B_n is convex, the matrix $D\nu = (D_j \nu_i)$ is positive definite. Moreover, differentiating the equality $\frac{\partial u}{\partial \nu} = 0$, one easily verifies that

$$\sum_{i,j=1}^d \nu_j D_{ij} u D_i u = - \sum_{i,j=1}^d D_i \nu_j D_i u D_j u \leq 0,$$

which, in its turn, implies that the normal derivative of z_n on ∂B_n is nonpositive.

We claim that we may choose $a > 0$ in such a way that $D_t z_n - \mathcal{A}(t) z_n \leq 0$ for $s < t < T$. Then, the classical maximum principle yields $|z_n| \leq \|f\|_{C_b^1(\mathbb{R}^d)}^2$, i.e.,

$$u_n(t, x)^2 + a|\nabla u_n(t, x)|^2 \leq \|f\|_{C_b^1(\mathbb{R}^d)}^2 \quad (t, x) \in (s, T) \times B_n.$$

Letting $n \rightarrow +\infty$, statement (i) follows with $C_1 = a^{-\frac{1}{2}}$.

From now on we omit the subscript n as well as the dependence on t and x to simplify notation. To prove the claim, observe that

$$(4.1) \quad \begin{aligned} z_t - A(\cdot)z &= 2a\langle \nabla_x b \nabla_x u, \nabla_x u \rangle - 2\langle Q \nabla_x u, \nabla_x u \rangle - 2a \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle \\ &\quad + 2a \sum_{k=1}^d D_k u \cdot \text{Tr}(D_k Q \cdot D_x^2 u). \end{aligned}$$

Using Hypothesis 1.2(iii), we estimate the last term as follows,

$$\left| \sum_{k=1}^d D_k u \cdot \text{Tr}(D_k Q D_x^2 u) \right| \leq \rho_0 \eta \sum_{k=1}^d |D_k u| \cdot \sum_{i,j=1}^d |D_{ij} u| \leq \rho_0 \eta d^{\frac{3}{2}} |\nabla_x u| |D_x^2 u|.$$

The other terms are easily estimated using Hypotheses 1.1(ii) and 1.2(ii). Eventually, we get

$$\begin{aligned} z_t - A(\cdot)z &\leq 2(ak_0 - \eta) |\nabla_x u|^2 - 2a\eta |D_x^2 u|^2 + 2a\rho_0 \eta d^{\frac{3}{2}} |\nabla_x u| |D_x^2 u| \\ &\leq 2(ak_0 - \eta) |\nabla_x u|^2 - 2a\eta |D_x^2 u|^2 + a\eta (\rho_0^2 d^3 |\nabla_x u|^2 + |D_x^2 u|^2) \\ &\leq (2ak_0 - 2\eta + a\eta \rho_0^2 d^3) |\nabla_x u|^2. \end{aligned}$$

The right hand side is negative, if we choose $a \leq d^{-3} \rho_0^{-2}$ such that $2ak_0 \leq \eta_0$.

(ii). We proceed similarly to (i), defining

$$z_n(t, x) = u_n(t, x)^2 + a(t-s)|\nabla_x u_n|^2, \quad (t, x) \in [s, T] \times B_n.$$

As above, in what follows we omit the subscript n as well as the dependence on t and x .

If we proceed as in part (i), we see that z satisfies an equality similar to (4.1) with a replaced by $a(t-s)$ and a further addendum $a|\nabla_x u|^2$. Hence,

$$z_t - A(\cdot)z \leq (2a(T-s)k_0^+ - 2\eta + a(T-s)\eta d^3 \rho_0^2 + a) |\nabla_x u|^2,$$

where $k_0^+ = \max\{k_0, 0\}$. The right-hand side is nonpositive, if we choose $a = a_T \leq (T-s)^{-1} d^{-3} \rho_0^{-2}$ such that $2a(T-s)k_0 + a \leq \eta_0$. By the maximum principle we obtain $z_n \leq \|f\|_\infty^2$ and statement (ii) follows, with $C_2 = a^{-\frac{1}{2}}$, letting $n \rightarrow +\infty$. \square

Remark 4.2. In the proof of Theorem 4.1 we have chosen to approximate $G(t, s)f$ by solutions of Cauchy-Neumann problems instead of Cauchy-Dirichlet problems as in the first part of the paper. Approximation by Cauchy-Dirichlet problems is in fact possible, but it requires stronger conditions on the coefficients (see e.g., [3, Section 6.1] for the autonomous case), that we want to avoid here.

As a consequence of Theorem 4.1, our evolution family enjoys the strong Feller property.

Corollary 4.3. *For any $f \in B_b(\mathbb{R}^d)$ and any $t > s$ we have $G(t, s)f \in C_b(\mathbb{R}^d)$.*

Proof. Let $f \in B_b(\mathbb{R}^d)$. Then, there exists a bounded sequence $(f_n) \subset C_b(\mathbb{R}^d)$ which converges pointwise to f almost everywhere in \mathbb{R}^d . As a consequence of Theorem 4.1, for any fixed $s < t$, the function $t \mapsto G(t, s)f_n$ is Lipschitz continuous with Lipschitz constant independent of n . The statement follows, observing that, by the dominated convergence theorem and (2.8), $G(t, s)f_n$ converges to $G(t, s)f$ pointwise. \square

Corollary 4.4. *For any $f \in C_b^1(\mathbb{R}^d)$ and $s \in I$, the function $\nabla G(\cdot, s)f$ is continuous in $[s, +\infty) \times \mathbb{R}^d$.*

Proof. We have to show only continuity at $t = s$. For any $n \in \mathbb{N}$, let $\varphi \in C_c^\infty(B_n)$ be such that $0 \leq \varphi \leq 1$ and $\varphi \equiv 1$ in B_{n-1} . Put $u(t, x) := (G(t, s)f)(x)$ and $v = \varphi u$. We have $v_t - A(t)v = \psi$ in B_n , where

$$\psi = -uA(t)\varphi - 2\langle Q\nabla\varphi, \nabla_x u \rangle.$$

From Theorems 2.2 and 4.1, it follows that the functions u and $\nabla_x u$ are bounded and continuous in $(s, T] \times \mathbb{R}^d$, for any $T > s$. Since φ is compactly supported in B_n , $\psi \in C((s, s+1], C_0(B_n))$. Moreover, Theorem 4.1 yields that $\|\psi\|_\infty \leq C\|f\|_{C_b^1(\mathbb{R}^d)}$ for some $C > 0$.

Let $\{G_n(t, s)\}$ be the evolution family associated with problem (2.2). By the variation of constants formula (e.g., [1, Proposition 3.2]) we have

$$(4.2) \quad v(t, \cdot) = G_n(t, s)(\varphi f) + \int_s^t G_n(t, \sigma)\psi(\sigma) d\sigma, \quad s < t < s+1.$$

By classical gradient estimates ([15, Chapter IV, Theorem 17]), we get

$$\|\nabla_x G_n(t, \sigma)\psi(\sigma)\|_\infty \leq \frac{C_1}{\sqrt{t-\sigma}}\|\psi(\sigma)\|_\infty \leq \frac{C_2}{\sqrt{t-\sigma}}\|f\|_{C_b^1(\mathbb{R}^d)},$$

for any $s < \sigma < t \leq s+1$ and some positive constants C_1 and C_2 . Hence, we can differentiate (4.2) obtaining

$$\nabla_x v(t) = \nabla_x G_n(t, s)(\varphi f) + \int_s^t \nabla_x G_n(t, \sigma)\psi(\sigma) d\sigma, \quad s < t < s+1.$$

Therefore, for any $x, x_0 \in B_{n-1}$ we have

$$|\nabla_x u(t, x) - \nabla f(x_0)| \leq |(\nabla_x G_n(t, s)(\varphi f))(x) - \nabla f(x_0)| + 2C_2\|f\|_{C_b^1(\mathbb{R}^d)}(t-s)^{\frac{1}{2}},$$

and this implies that $\nabla G(\cdot, s)f$ is continuous at the point (s, x_0) since the function $\nabla_x G_n(\cdot, s)(\varphi f)$ is continuous in $\{s\} \times \overline{B_n}$ by classical results. Since n is arbitrary, the statement follows. \square

Next, we prove a pointwise gradient estimate.

Theorem 4.5. *Assume that Hypotheses 1.1, 1.2(i)(iii) and 1.3(i) are satisfied. Then for every $p \geq p_0$ and any $f \in C_b^1(\mathbb{R}^d)$ we have*

$$(4.3) \quad |(\nabla G(t, s)f)(x)|^p \leq e^{\sigma_p(t-s)}(G(t, s)|\nabla f|^p)(x), \quad t \geq s, \quad x \in \mathbb{R}^d,$$

where

$$(4.4) \quad \sigma_p = p \sup_{(t,x) \in I \times \mathbb{R}^d} \left(r(t,x) + \frac{d^3(\rho(t))^2 \eta(t,x)}{4 \min\{p_0 - 1, 1\}} \right).$$

Similarly, under Hypotheses 1.1, 1.2(i) and 1.3(ii), estimate (4.3) holds true for any $p \in (1, +\infty)$, with

$$(4.5) \quad \sigma_p = p \left(k_0 + \frac{d^3 \rho_0^2}{4 \min\{p - 1, 1\}} \right).$$

Moreover, if the coefficients q_{ij} ($i, j = 1, \dots, d$) do not depend on x and $r \leq \eta$ in $I \times \mathbb{R}^d$, then (4.5) holds true for $p = 1$ too, provided Hypothesis 1.3(ii) is satisfied. In such a case, $\sigma_1 = k_0$.

Proof. To prove the first part of the statement, fix $s \in I$ and $\varepsilon > 0$. Set $u(t, x) := (G(t, s)f)(x)$ and define

$$w(t, x) = (|\nabla_x u(t, x)|^2 + \varepsilon)^{\frac{p}{2}}, \quad t \geq s, \quad x \in \mathbb{R}^d.$$

By Corollary 4.4, $w \in C_b([s, T] \times \mathbb{R}^d)$ for all $T > s$, and moreover by [9, Theorem 3.10], $w \in C^{1,2}((s, T) \times \mathbb{R}^d)$. A straightforward computation shows that

$$w_t - \mathcal{A}(t)w = f_1 + f_2 + f_3,$$

where

$$\begin{aligned} f_1 &= p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \left(\langle \nabla_x b \nabla_x u, \nabla_x u \rangle - \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle \right), \\ f_2 &= p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \sum_{k=1}^d D_k u \cdot \text{Tr}(D_k Q \cdot D_x^2 u), \\ f_3 &= -p(p-2)(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-2} \langle Q D_x^2 u \nabla_x u, D_x^2 u \nabla_x u \rangle. \end{aligned}$$

Using Hypotheses 1.1(ii), 1.2(i)(iii) and 1.3(i), we estimate f_1 as in the proof of Theorem 4.1, getting

$$\begin{aligned} (4.6) \quad f_1 &\leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \cdot \left(r |\nabla_x u|^2 - \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle \right) \\ &\leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \cdot (r |\nabla_x u|^2 - \eta |D_x^2 u|^2). \end{aligned}$$

Moreover, for every $c > 0$ we have

$$(4.7) \quad f_2 \leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \eta \left(c |D_x^2 u|^2 + \frac{d^3 \rho^2}{4c} |\nabla_x u|^2 \right).$$

Concerning f_3 , we have

$$\begin{aligned} (4.8) \quad \langle Q D_x^2 u \nabla_x u, D_x^2 u \nabla_x u \rangle &= |Q^{1/2} D_x^2 u \nabla_x u|^2 \\ &\leq \|Q^{1/2} D_x^2 u\|^2 |\nabla_x u|^2 \\ &= |\nabla_x u|^2 \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle. \end{aligned}$$

Now we distinguish between two cases.

Case 1: $p \geq \max\{p_0, 2\}$. Since $p(p-2) \geq 0$, the uniform ellipticity assumption implies $f_3 \leq 0$. Using (4.6) and (4.7) with $c = 1$, we obtain

$$\begin{aligned} (4.9) \quad w_t - A(\cdot)w &\leq \sigma_p (|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} |\nabla_x u|^2 \\ &= \sigma_p (|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}} - \sigma_p \varepsilon (|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1}. \end{aligned}$$

Now, observing that

$$a^{\frac{p}{2}-1} \leq \left(1 - \frac{2}{p}\right) a^{\frac{p}{2}} + \frac{p}{2}, \quad a > 0,$$

from (4.9) we deduce that

$$w_t - A(\cdot)w \leq \sigma_{p,\varepsilon}(w - \delta_\varepsilon),$$

where,

$$\sigma_{p,\varepsilon} = \begin{cases} \sigma_p, & \text{if } \sigma_p \geq 0, \\ \sigma_p \left\{1 - \left(1 - \frac{2}{p}\right) \varepsilon\right\}, & \text{if } \sigma_p < 0, \end{cases} \quad \delta_\varepsilon = \begin{cases} 0, & \text{if } \sigma_p \geq 0, \\ \frac{p}{2}\varepsilon, & \text{if } \sigma_p < 0. \end{cases}$$

and $\sigma_{p,\varepsilon}$ is given by (4.4).

Case 2: $p_0 < 2$ and $1 < p < 2$. In this case, $-p(p-2)$ is positive. Hence, we may use (4.8) to estimate f_3 . Together with estimates (4.7) and (4.8) (with $c = p-1$), we obtain

$$\begin{aligned} w_t - A(\cdot)w &\leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \\ &\quad \times \left((2-p) \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle + \frac{d^3 \rho^2}{4(p-1)} \eta |\nabla_x u|^2 \right. \\ &\quad \left. + r |\nabla_x u|^2 - \eta \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle + (p-1)\eta |D_x^2 u|^2 \right) \\ &\leq \sigma_{p,\varepsilon}(w - \delta_\varepsilon). \end{aligned}$$

Here,

$$\delta_\varepsilon = \begin{cases} 0, & \text{if } \sigma_p \geq 0, \\ \varepsilon^{\frac{p}{2}}, & \text{if } \sigma_p < 0, \end{cases}$$

and $\sigma_{p,\varepsilon} := \sigma_p$ is given by (4.4).

Now the procedure is the same in the two cases. Setting $v = w - \delta_\varepsilon$ we have $v_t - A(\cdot)v \leq \sigma_{p,\varepsilon}v$. On the other hand, the function

$$z(t) = e^{\sigma_{p,\varepsilon}(t-s)} G(t, s) (|\nabla f|^2 + \varepsilon)^{\frac{p}{2}}, \quad t > s,$$

satisfies $z_t - \mathcal{A}(t)z = \sigma_{p,\varepsilon}z$. Thus,

$$\begin{cases} (v-z)_t - (\mathcal{A}(t) + \sigma_{p,\varepsilon})(v-z) \leq 0, & t \in (s, +\infty), \\ (v-z)(s) = -\delta_\varepsilon. \end{cases}$$

Theorem 2.1 implies $v \leq z$. Letting $\varepsilon \rightarrow 0^+$, the statement follows by Proposition 3.1.

In the case that Hypothesis (1.3)(i) is replaced by (1.3)(ii), the functions f_1, f_2 are estimated as follows:

$$(4.10) \quad f_1 \leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \cdot \left(k_0 |\nabla_x u|^2 - \sum_{k=1}^d \langle Q \nabla_x D_k u, \nabla_x D_k u \rangle \right)$$

$$(4.11) \quad \leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \cdot (k_0 |\nabla_x u|^2 - \eta |D_x^2 u|^2),$$

$$(4.12) \quad f_2 \leq p(|\nabla_x u|^2 + \varepsilon)^{\frac{p}{2}-1} \left(c\eta |D_x^2 u|^2 + \frac{d^3 \rho_0^2}{4c} |\nabla_x u|^2 \right),$$

for any $c > 0$. Then, estimate (4.3) with $p \in (1, +\infty)$ (and with $p = 1$ too, if the diffusion coefficients are constant with respect to x), follows arguing as above. \square

Corollary 4.6. *Under the hypotheses of Theorem 4.5, there exists a constant C such that*

$$\|\nabla G(t, s)f\|_\infty \leq C \cdot e^{\frac{\sigma_p}{p}(t-s)} \|f\|_\infty, \quad f \in C_b(\mathbb{R}^d), \quad s \in I, \quad t \geq s+1,$$

for every $p \geq p_0$ if Hypothesis 1.3(i) is satisfied, and for every $p > 1$ if Hypothesis 1.3(ii) is satisfied.

Proof. By Theorem 4.1, for any $f \in C_b(\mathbb{R}^d)$ the function $G(s+1, s)f$ is in $C_b^1(\mathbb{R}^d)$, and its C^1 -norm does not exceed $C_1\|f\|_\infty$ for some $C_1 > 0$, independent of f . If $t > s+1$, we have, by Theorem 4.5,

$$\begin{aligned} |(\nabla G(t, s)f)(x)|^p &= |(\nabla G(t, s+1)G(s+1, s)f)(x)| \\ &\leq e^{\sigma_p(t-(s+1))} (G(t, s+1)|\nabla G(s+1, s)f|^p)(x) \\ &\leq e^{\sigma_p(t-(s+1))} \|\nabla G(s+1, s)f\|_\infty^p. \end{aligned}$$

Thus,

$$\|\nabla G(t, s)f\|_\infty^p \leq e^{\sigma_p(t-(s+1))} \|\nabla G(s+1, s)f\|_\infty^p,$$

and the statement follows. \square

5. EVOLUTION SYSTEMS OF MEASURES

Definition 5.1. Let $\{U(t, s)\}$ be an evolution family of bounded operators on $B_b(\mathbb{R}^d)$. A family (ν_t) of probability measures on \mathbb{R}^d is an *evolution system of measures* for $\{U(t, s)\}$ if, for every $f \in B_b(\mathbb{R}^d)$ and every $s < t$, we have

$$(5.1) \quad \int_{\mathbb{R}^d} U(t, s)f \, d\nu_t = \int_{\mathbb{R}^d} f \, d\nu_s.$$

Formula (5.1) may be rewritten as $U^*(t, s)\nu_t = \nu_s$. It implies that, if we know a single measure ν_{t_0} of an evolution system of measures for $\{U(t, s)\}$, then we know all the measures ν_t for $t \leq t_0$. In particular, an evolution system of measures is uniquely determined by its tail $(\nu_t)_{t \geq t_0}$.

In this section we give sufficient conditions for the existence of an evolution system (μ_t) of measures associated with the evolution family $\{G(t, s)\}$ and we study the main properties of (μ_t) . As a first step, we note that, for our evolution family $\{G(t, s)\}$, evolution systems of measures necessarily consist of measures which are equivalent to the Lebesgue measure.

Proposition 5.2. *If (μ_t) is an evolution system of measures for $\{G(t, s)\}$ then (μ_t) is equivalent to the Lebesgue measure.*

Proof. For each $A \in B(\mathbb{R}^d)$ and $t \in I$ we have

$$\mu_t(A) = \int_{\mathbb{R}^d} (G(t+1, t)\mathbf{1}_A)(x) \, \mu_{t+1}(dx).$$

By Corollaries 2.5 and 4.3, if the Lebesgue measure $|A|$ of A is positive then $(G(t+1, t)\mathbf{1}_A)(x)$ is positive for each $x \in \mathbb{R}^d$; therefore $\mu_t(A) > 0$. On the other hand, by Proposition 2.4(ii), if $|A| = 0$ then $G(t+1, t)\mathbf{1}_A = 0$, hence $\mu_t(A) = 0$. \square

To prove existence of evolution systems of measures we use a procedure similar to the Krylov-Bogoliubov Theorem which states that, in the autonomous case, existence of an invariant measure is equivalent to the tightness of a certain set of probability measures. In our case, the corresponding tightness property is proved under Hypothesis 1.4, through the Prokhorov Theorem. It states that a set $\{P_\alpha : \alpha \in \mathcal{F}\}$ of probability measures is tight if and only if, for any sequence (α_n) in

\mathcal{F} , there exists a subsequence α_{n_k} such that $P_{\alpha_{n_k}}$ converges to some probability measure P in the following sense:

$$\lim_{k \rightarrow +\infty} \int_{\mathbb{R}^d} f(y) P_{\alpha_k}(dy) = \int_{\mathbb{R}^d} f(y) P(dy), \quad f \in C_b(\mathbb{R}^d).$$

Lemma 5.3. *Assume that Hypotheses 1.1 and 1.4 are satisfied. Then, $G(t, s)\varphi$ is well defined for any $t_0 \leq s \leq t \in I$. Moreover, for any fixed $x \in \mathbb{R}^d$, the function $(t, s) \mapsto (G(t, s)\varphi)(x)$ is bounded in $\Lambda = \{(t, s) \in I \times I : t_0 \leq s \leq t\}$.*

Proof. Lemma 3.4 implies that $G(t, s)\varphi$ is well defined for $(t, s) \in \Lambda$ with $t_0 \leq s$ and the function $(t, s, x) \mapsto (G(t, s)\varphi)(x)$ is locally bounded. To complete the proof, we fix $t > t_0$ and $x \in \mathbb{R}^d$, and consider the function g defined in $[t_0, t]$ by $g(s) := (G(t, s)\varphi)(x)$. g is measurable, because $(G(t, s)\varphi)(x)$ is the pointwise limit of the functions $(G(t, s)\varphi_n)(x)$ in the proof of Lemma 3.4, that are continuous with respect to s . The procedure of Lemma 3.4 yields

$$(5.2) \quad g(r) - g(s) \geq \int_s^r (cg(\sigma) - a) d\sigma, \quad t_0 \leq s \leq r \leq t.$$

We claim that (5.2) implies

$$(5.3) \quad g(s) \leq \left(g(t) - \frac{a}{c} \right) e^{c(s-t)} + \frac{a}{c}, \quad t_0 \leq s \leq t.$$

Indeed, for any fixed $s \geq t_0$, the function Φ defined by

$$\Phi(r) := \left(g(s) - \frac{a}{c} + \int_s^r (cg(\sigma) - a) d\sigma \right) e^{-cr}, \quad r \leq s \leq t$$

is continuous in $[s, t]$ and therein weakly differentiable with $\Phi'(r) \geq 0$ a.e., so that it is nondecreasing, and $\Phi(s) \leq \Phi(t)$ implies (5.3). From (5.3) we obtain $(G(t, s)\varphi)(x) \leq \varphi(x) + a/c$, and the statement follows. \square

Theorem 5.4. *Assume that Hypotheses 1.1 and 1.4 are satisfied. Then, there exist an evolution system (μ_t) of measures for $\{G(t, s)\}$ and a constant $M \geq 0$ such that*

$$(5.4) \quad \int_{\mathbb{R}^d} \varphi(y) \mu_t(dy) \leq M, \quad t \geq t_0.$$

Proof. Fix $s \in I$ and $x_0 \in \mathbb{R}^d$. For any $t > s$, define the measure $\mu_{t,s}$ by

$$\mu_{t,s}(A) := \frac{1}{t-s} \int_s^t p_{\tau,s}(x_0, A) d\tau = \frac{1}{t-s} \int_s^t (G(\tau, s)\mathbf{1}_A)(x_0) d\tau.$$

Lemma 5.3 implies that the family $(\mu_{t,s})_{t>s \geq t_0}$ is tight, through the same proof of Lemma 3.5. The Prokhorov Theorem and a diagonal argument yield existence of a sequence t_k diverging to $+\infty$ and of probability measures μ_n ($n \in \mathbb{N}$, $n > t_0$) such that $\mu_{t_k,n} \rightharpoonup^* \mu_n$. To define μ_s also for noninteger s , we show preliminarily that

$G^*(n, m)\mu_n = \mu_m$ for $m < n$. Indeed, for each $A \in \mathcal{B}(\mathbb{R}^d)$ we have

$$\begin{aligned} G^*(n, m)\mu_n(A) &= \int_{\mathbb{R}^d} \mathbb{1}_A(y) G^*(n, m)\mu_n(dy) \\ &= \int_{\mathbb{R}^d} (G(n, m)\mathbb{1}_A)(y) \mu_n(dy) \\ &= \lim_{k \rightarrow +\infty} \frac{1}{t_k - n} \int_n^{t_k} (G(\tau, n)G(n, m)\mathbb{1}_A)(x) d\tau \\ &= \lim_{k \rightarrow +\infty} \frac{1}{t_k - n} \int_n^{t_k} (G(\tau, m)\mathbb{1}_A)(x) d\tau \\ &= \lim_{k \rightarrow +\infty} \frac{1}{t_k - m} \int_m^{t_k} (G(\tau, m)\mathbb{1}_A)(x) d\tau \\ &= \mu_m(A). \end{aligned}$$

Thus, we can extend the definition of the measures μ_s to any $s \in I$, by setting $\mu_s := G^*(n, s)\mu_n$ where n is any positive integer greater than s . Since $G^*(n, s) = G^*(m, s)G^*(n, m)$, this definition is independent of n . It is immediate to check that (μ_t) is an evolution system of measures for $\{G(t, s)\}$.

To complete the proof, we observe that, since $(G(t, s)\varphi)(x_0)$ is bounded in $t \geq s \geq t_0$, then each integral

$$\frac{1}{t-s} \int_s^t (G(\tau, s)\varphi)(x_0) d\tau = \int_{\mathbb{R}^d} \varphi(y) \mu_{t,s}(dy)$$

is bounded for $t > s \geq t_0$ by the same constant. Letting $t \rightarrow +\infty$, we get (5.4). \square

Remark 5.5. It should be noted that the evolution system of measures constructed in Theorem 5.4 could still depend on x_0 . Indeed, in general, evolution system of measures are not unique. In [10, Lemma 2.2] it is proved that the evolution family associated with the operators

$$(\mathcal{A}(t)u)(x) = \frac{1}{2} \Delta u(x) + \langle B(t)x, \nabla u(x) \rangle,$$

admits infinitely many evolution systems of measures. However, uniqueness may be achieved among all systems of measures which have finite moments of order p for some $p > 0$ with a certain asymptotic behaviour.

In the following, if (μ_t) is a family of probability measures on \mathbb{R}^d , we denote by

$$\mu_t(p) := \int_{\mathbb{R}^d} |x|^p \mu_t(dx),$$

the p -th moment function. We note that, if $\varphi(x) = |x|^p$ satisfies Hypothesis 1.4, then Theorem 5.4 implies that $\{G(t, s)\}$ admits an evolution system of measures (μ_t) such that $\mu_t(p) = O(1)$ as $t \rightarrow +\infty$, i.e. there exists $t_0 \in I$ such that the p -th moments of μ_t exist and are uniformly bounded for any $t \geq t_0$.

Let us see the connection between evolution systems of measures and asymptotic behaviour of solutions to problem (1.2). We assume that there exists a *negative* constant ω such that, for large $t - s$, we have

$$\|\nabla G(t, s)f\|_\infty \leq e^{\omega(t-s)} \|f\|_\infty, \quad f \in C_b(\mathbb{R}^d).$$

A sufficient condition for this may be obtained from Corollary 4.6.

Theorem 5.6. *Assume that there exists $\omega < 0$ such that*

$$(5.5) \quad \|\nabla G(t, s)f\|_\infty \leq C e^{\omega(t-s)} \|f\|_\infty,$$

for all $t \geq s+1$, all $f \in C_b(\mathbb{R}^d)$ and some positive constant C . Further, assume that $\{G(t, s)\}$ admits an evolution system of measures (μ_t) such that, for some $p > 0$,

$$\lim_{t \rightarrow +\infty} \mu_t(p) e^{\omega p t} = 0.$$

Then,

$$\lim_{t \rightarrow +\infty} (G(t, s)f)(x) = \int_{\mathbb{R}^d} f(y) \mu_s(dy), \quad x \in \mathbb{R}^d,$$

for all $s \in I$ and $f \in C_b(\mathbb{R}^d)$. If $I = \mathbb{R}$, then, we also have

$$\lim_{s \rightarrow -\infty} \left((G(t, s)f)(x) - \int_{\mathbb{R}^d} f(y) \mu_s(dy) \right) = 0, \quad x \in \mathbb{R}^d.$$

In both cases the convergence is uniform in the compact sets of \mathbb{R}^d .

Proof. Without loss of generality, we may assume that $p < 1$. We have

$$\begin{aligned} (G(t, s)f)(x) - \int_{\mathbb{R}^d} f(y) \mu_s(dy) &= (G(t, s)f)(x) - \int_{\mathbb{R}^d} (G(t, s)f)(y) \mu_t(dy) \\ &= \int_{\mathbb{R}^d} \{(G(t, s)f)(x) - (G(t, s)f)(y)\} \mu_t(dy). \end{aligned}$$

Splitting

$$\begin{aligned} &|(G(t, s)f)(x) - (G(t, s)f)(y)| \\ &= |(G(t, s)f)(x) - (G(t, s)f)(y)|^{1-p} |(G(t, s)f)(x) - (G(t, s)f)(y)|^p \end{aligned}$$

and using the mean value theorem and (5.5), we get

$$|(G(t, s)f)(x) - (G(t, s)f)(y)| \leq 2C^p \|f\|_\infty e^{p\omega(t-s)} |x - y|^p.$$

Hence, we have:

$$\begin{aligned} \left| (G(t, s)f)(x) - \int_{\mathbb{R}^d} f(y) \mu_s(dy) \right| &\leq 2C^p \|f\|_\infty e^{p\omega(t-s)} \int_{\mathbb{R}^d} |x - y|^p \mu_t(dy) \\ (5.6) \quad &\leq 2C^p \|f\|_\infty e^{p\omega(t-s)} \left(|x|^p + \int_{\mathbb{R}^d} |y|^p \mu_t(dy) \right), \end{aligned}$$

and the right-hand side vanishes as $t \rightarrow +\infty$ (and also as $s \rightarrow -\infty$, if $I = \mathbb{R}$), uniformly for x in compact sets. \square

Corollary 5.7. *Under the hypothesis of Theorem 5.6, there exists at most one evolution system of measures (μ_t) such that $\lim_{t \rightarrow +\infty} \mu_t(p) e^{\omega p t} = 0$ for some $p > 0$.*

Proof. Let (μ_t) , (ν_t) be two evolution system of measures with the above property. By Theorem 5.6, for each $f \in C_b(\mathbb{R}^d)$ and $s \in I$ we have

$$\int_{\mathbb{R}^d} f(y) \mu_s(dy) = \int_{\mathbb{R}^d} f(y) \nu_s(dy),$$

since both integrals coincide with $\lim_{t \rightarrow +\infty} (G(t, s)f)(0)$. The statement follows. \square

6. EVOLUTION SEMIGROUPS IN L^p SPACES WITH RESPECT TO INVARIANT MEASURES

In this section we assume that $I = \mathbb{R}$, and that Hypotheses 1.1 and 1.4 are satisfied.

Let us define the evolution semigroup $\{\mathcal{T}(t)\}$ associated with the evolution family $\{G(t, s)\}$ on the space $C_b(\mathbb{R}^{d+1})$ by

$$(\mathcal{T}(t)f)(s, x) = (G(s, s-t)f(s-t, \cdot))(x), \quad (s, x) \in \mathbb{R}^{d+1}, \quad t \geq 0.$$

Proposition 6.1. *The family of operators $\{\mathcal{T}(t) : t \geq 0\}$ is a semigroup of positive contractions in $C_b(\mathbb{R}^{d+1})$. Moreover, $\mathcal{T}(t)f$ tends to f locally uniformly in \mathbb{R}^{d+1} as $t \rightarrow 0^+$, for any $f \in C_b(\mathbb{R}^{d+1})$.*

Proof. As a first step we prove that, for any $t > 0$, the operator $\mathcal{T}(t)$ maps $C_b(\mathbb{R}^{d+1})$ into itself. By Theorem 2.2, we know that

$$\sup_{(s,x) \in \mathbb{R}^{d+1}} |(\mathcal{T}(t)f)(s,x)| \leq \|f\|_{C_b(\mathbb{R}^{d+1})}, \quad t \geq 0.$$

Let us now fix (s_0, x_0) in \mathbb{R}^{d+1} and observe that

$$\begin{aligned} & |(\mathcal{T}(t)f)(s,x) - (\mathcal{T}(t)f)(s_0,x_0)| \\ &= |(G(s,s-t)f(s,\cdot))(x) - (G(s_0,s_0-t)f(s_0,\cdot))(x_0)| \\ &\leq |(G(s,s-t)f(s,\cdot))(x) - (G(s,s-t)f(s_0,\cdot))(x)| \\ (6.1) \quad & + |(G(s,s-t)f(s_0,\cdot))(x) - (G(s_0,s_0-t)f(s_0,\cdot))(x_0)|. \end{aligned}$$

By Proposition 3.6,

$$\lim_{s \rightarrow s_0} \sup_{(r,x) \in [s_0-\delta, s_0+\delta] \times \{x_0+B(\delta)\}} |(G(r,r-t)f(s,\cdot))(x) - (G(r,r-t)f(s_0,\cdot))(x)| = 0,$$

for any $\delta > 0$. Therefore, the first term in the right-hand side of (6.1) converges to 0 as (s,x) tends to (s_0,x_0) . Similarly, by Theorem 3.7, the function $(p,r,x) \mapsto (G(p,r)f)(x)$ is continuous in $\{(p,r,x) \in \mathbb{R}^{d+2} : r \leq p\}$. Hence, also the second term tends to 0 as (s,x) tends to (s_0,x_0) . This shows that $\mathcal{T}(t)f \in C_b(\mathbb{R}^{d+1})$.

The semigroup property follows easily since $\{G(t,s)\}$ is an evolution family. Indeed, for any $t_1 < t_2$, it holds that

$$\begin{aligned} & (\mathcal{T}(t_2)\mathcal{T}(t_1)f)(s,x) \\ &= (G(s,s-t_2)\mathcal{T}(t_1)f(s-t_2,\cdot))(x) \\ &= (G(s,s-t_2)G(s-t_2,s-t_2-t_1)f(s-t_1-t_2,\cdot))(x) \\ &= (G(s,s-t_2-t_1)f(s-t_1-t_2,\cdot))(x) \\ &= (\mathcal{T}(t_1+t_2)f)(s,x), \end{aligned}$$

for any $(s,x) \in \mathbb{R}^{d+1}$.

The positivity of $\mathcal{T}(t)$ follows from the positivity of the evolution family $\{G(t,s)\}$.

Finally, the fact that $\mathcal{T}(t)f$ converges to f locally uniformly in \mathbb{R}^{d+1} as $t \rightarrow 0^+$, is an immediate consequence of the continuity of the function $(p,r,x) \mapsto (G(p,r)f)(x)$ in $\{(p,r,x) \in \mathbb{R}^{d+2} : p \geq r\}$ and Proposition 3.6. \square

Remark 6.2. Since $G(s,s-t)\mathbf{1} = \mathbf{1}$ for each $t > s$, if $f = f(s)$ depends only on time then $(\mathcal{T}(t)f)(s,x) = f(s-t)$, i.e. $\mathcal{T}(t)$ acts as a translation semigroup. Therefore, $\mathcal{T}(t)$ cannot have any smoothing or summability improving property in the s variable. In particular, it is not strong Feller and not hypercontractive.

Now, let (μ_t) be an evolution system of measures for $\{G(t,s)\}$. Note that the function $s \mapsto \mu_s(A)$ is measurable in I for any Borel set A . Indeed, by Lemma 3.2, the function $s \mapsto (G(t,s)f)(x)$ is bounded and continuous in $(-\infty, t)$, for any $x \in \mathbb{R}^d$ and any $f \in C_0(\mathbb{R}^d)$. Hence, the function

$$s \mapsto \int_{\mathbb{R}^d} (G(t,s)f)(x) \mu_t(dx),$$

is continuous as well in $(-\infty, t)$. Since

$$\mu_s(A) = \int_{\mathbb{R}^d} (G(t,s)\mathbf{1}_A)(x) \mu_t(dx),$$

and $\mathbb{1}_A$ is the pointwise limit of a sequence $(f_n) \subset C_0(\mathbb{R}^d)$, bounded with respect to the sup-norm, by dominated convergence, the measurability of the function $s \mapsto \mu_s(A)$ follows. Therefore, we can define

$$\nu(J \times K) := \int_J \mu_t(K) dt,$$

for Borel sets $J \subset \mathbb{R}$ and $K \subset \mathbb{R}^d$. Of course, ν may be uniquely extended in a standard way to a measure on $\mathcal{B}(\mathbb{R}^{d+1})$.

In the following, we denote by \mathcal{G} the differential operator

$$(6.2) \quad \mathcal{G}u(t, x) = \mathcal{A}(t)u(t, x) - u_t(t, x), \quad (t, x) \in \mathbb{R}^{d+1}.$$

We state a preliminary lemma about $\mathcal{T}(t)$ and \mathcal{G} .

Lemma 6.3. (i) For all $\varphi \in C_c(\mathbb{R}, C_b(\mathbb{R}^d))$ and all $t \geq 0$ we have

$$\int_{\mathbb{R}^{d+1}} \mathcal{T}(t)\varphi d\nu = \int_{\mathbb{R}^{d+1}} \varphi d\nu.$$

(ii) For $\varphi \in C_c^{1,2}(\mathbb{R}^{d+1})$ we have

$$(6.3) \quad \int_{\mathbb{R}^{d+1}} \mathcal{G}\varphi d\nu = 0.$$

Proof. (i). We have

$$\begin{aligned} \int_{\mathbb{R}^{d+1}} \mathcal{T}(t)\varphi d\nu &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} (G(s, s-t)\varphi(s-t, \cdot))(x) \mu_s(dx) ds \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} \varphi(s-t, x) \mu_{s-t}(dx) ds \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} \varphi(r, x) \mu_r(dx) dr \\ &= \int_{\mathbb{R}^{d+1}} \varphi d\nu. \end{aligned}$$

(ii). By part (i) we obtain

$$\lim_{h \rightarrow 0^+} \int_{\mathbb{R}^{d+1}} \frac{\mathcal{T}(h)\varphi - \varphi}{h} d\nu = 0.$$

Now we show that

$$\lim_{h \rightarrow 0^+} \int_{\mathbb{R}^{d+1}} \frac{\mathcal{T}(h)\varphi - \varphi}{h} d\nu = \int_{\mathbb{R}^{d+1}} \mathcal{G}\varphi d\nu.$$

For this purpose, let $a, b \in \mathbb{R}$ and $\delta > 0$ be such that $\text{supp}(\varphi) \subset [a, b] \times B(\delta)$. Then, if $t \in [0, 1]$, the support of the function $(s, x) \mapsto \varphi(s-t, x)$ is contained in

$[a, b+1] \times B(\delta)$. Therefore, for any $h \in (0, 1]$ we have

$$\begin{aligned} & \int_{\mathbb{R}^{d+1}} \frac{\mathcal{T}(h)\varphi - \varphi}{h} d\nu \\ &= \int_{[a, b+1] \times \mathbb{R}^d} \frac{\mathcal{T}(h)\varphi - \varphi}{h} d\nu \\ &= \int_a^{b+1} \int_{\mathbb{R}^d} \frac{(G(s, s-h)\varphi(s-h, \cdot))(x) - (G(s, s-h)\varphi(s, \cdot))(x)}{h} \mu_s(dx) ds \\ &\quad + \int_a^{b+1} \int_{\mathbb{R}^d} \frac{(G(s, s-h)\varphi(s, \cdot))(x) - \varphi(s, x)}{h} \mu_s(dx) ds \\ &=: \int_a^{b+1} (I_1(s, h) + I_2(s, h)) ds. \end{aligned}$$

As far as I_1 is concerned, we note that

$$\sup_{\mathbb{R}^d} \left| G(s, s-h) \left(\frac{\varphi(s-h, \cdot) - \varphi(s, \cdot)}{h} \right) \right| \leq \sup_{\mathbb{R}^{d+1}} |\varphi_t| < +\infty,$$

and moreover

$$\lim_{h \rightarrow 0^+} \frac{\varphi(s-h, x) - \varphi(s, x)}{h} = -\varphi_s(s, x),$$

the convergence being uniform in x . Since $f := -\varphi_s \in C_c(\mathbb{R}^{d+1})$, $G(s, s-h)f$ converges uniformly to f as $h \rightarrow 0^+$ by Lemma 3.2. Overall we see that

$$\lim_{h \rightarrow 0^+} I_1(s, h) = \int_{\mathbb{R}^d} -\varphi_s(s, x) \mu_s(dx),$$

and $I_1(s, h)$ is bounded by $\sup_{\mathbb{R}^d} |D_t \varphi|$.

Let us consider $I_2(s, h)$. Taking Lemma 3.2 into account, we write

$$\begin{aligned} I_2(s, h) &= \frac{1}{h} \int_{\mathbb{R}^d} \int_{s-h}^s (G(s, r)\mathcal{A}(r)\varphi(s, \cdot))(x) dr \mu_s(dx) \\ &= \frac{1}{h} \int_{s-h}^s \int_{\mathbb{R}^d} (G(s, r)\mathcal{A}(r)\varphi(s, \cdot))(x) \mu_s(dx) dr \\ &= \frac{1}{h} \int_{s-h}^s \int_{\mathbb{R}^d} \mathcal{A}(r)\varphi(s, x) \mu_r(dx) dr, \end{aligned}$$

so that

$$\lim_{h \rightarrow 0^+} I_2(s, h) = \int_{\mathbb{R}^d} \mathcal{A}(s)\varphi(s, x) \mu_s(dx),$$

for almost every s , by the Lebesgue differentiation theorem. We also note that

$$\sup_{s \in [a, b+1]} |I_2(s, h)| \leq \sup_{\substack{r \in [a-1, b+1] \\ (s, x) \in \text{supp}(\varphi)}} |(\mathcal{A}(r)\varphi(s, \cdot))(x)|.$$

Hence, by the dominated convergence theorem,

$$\lim_{h \rightarrow 0^+} \int_{\mathbb{R}} (I_1(s, h) + I_2(s, h)) ds = \int_{\mathbb{R}} \int_{\mathbb{R}^d} (-\varphi_s(s, x) + \mathcal{A}(s)\varphi(s, x)) \mu_s(dx) ds.$$

This proves (ii). \square

Remark 6.4. In view of (6.3) we say that ν is infinitesimally invariant, although it is not a probability measure.

Proposition 6.5. *For any $p \in [1, +\infty)$, the semigroup $\{\mathcal{T}(t)\}$ extends uniquely to a strongly continuous semigroup of positive contractions $\{\mathcal{T}_p(t)\}$ on $L^p(\mathbb{R}^{d+1}, \nu)$. Moreover, the infinitesimal generator of $\{\mathcal{T}_p(t)\}$ is an extension of the operator $\mathcal{G}_0 : C_c^{1,2}(\mathbb{R}^{d+1}) \rightarrow L^p(\mathbb{R}^{d+1}, \nu)$ defined by $\mathcal{G}_0 f = \mathcal{G}f$, for any $f \in C_c^{1,2}(\mathbb{R}^{d+1})$, where \mathcal{G} is given by (6.2).*

Proof. Using the Hölder inequality and taking proposition 2.4 into account, it is immediate to check that

$$|(\mathcal{T}(t)f)(s, x)|^p \leq (\mathcal{T}(t)|f|^p)(s, x), \quad (s, x) \in \mathbb{R}^{d+1}, \quad t > 0,$$

for any $f \in C_c(\mathbb{R}^{d+1})$. Integrating in \mathbb{R}^{d+1} , we obtain

$$(6.4) \quad \|\mathcal{T}(t)f\|_{L^p(\mathbb{R}^{d+1}, \nu)} \leq \|f\|_{L^p(\mathbb{R}^{d+1}, \nu)}, \quad t > 0.$$

Since $C_c(\mathbb{R}^{d+1})$ is dense in $L^p(\mathbb{R}^{d+1}, \nu)$, estimate (6.4) implies that any operator $\mathcal{T}(t)$ can be extended uniquely to a bounded operator $\mathcal{T}_p(t)$ which also satisfies (6.4).

Clearly, $\{\mathcal{T}_p(t)\}$ satisfies the semigroup law since $\{\mathcal{T}(t)\}$ does. It remains to show that $\{\mathcal{T}_p(t)\}$ is strongly continuous. Of course, it suffices to show that $\mathcal{T}_p(t)f \rightarrow f$ as $t \rightarrow 0^+$ for all $f \in C_c^{1,2}(\mathbb{R}^{d+1})$. For such f 's, we have $\mathcal{T}_p(t)f = \mathcal{T}(t)f \rightarrow f$ pointwise a.e. as $t \rightarrow 0^+$ (see the proof of Lemma 6.3(2), where it was shown that the difference quotients converge pointwise a.e.) and the functions $\mathcal{T}_p(t)f$ are uniformly bounded. The dominated convergence theorem implies that $\{\mathcal{T}_p(t)\}$ is strongly continuous.

To complete the proof, let us prove that $C_c^{1,2}(\mathbb{R}^{d+1})$ is contained in the domain of the infinitesimal generator of the semigroup $\{\mathcal{T}_p(t)\}$. For this purpose, we adapt the proof of Lemma 6.3(ii). Let $a, b \in \mathbb{R}$ and $\delta > 0$ be such that $\text{supp}(\varphi) \subset [a, b] \times B(\delta)$.

By Lemma 3.2 we know that

$$(G(s, s-t)f(s-t, \cdot))(x) - f(s-t, x) = \int_{s-t}^s (G(s, r)\mathcal{A}(r)f(s-t, \cdot))(x) dr,$$

for any $(s, x) \in \mathbb{R}^{d+1}$ and any $t > 0$. It follows that

$$\begin{aligned} & \left| \frac{(\mathcal{T}(t)f)(s, x) - f(s, x)}{t} - (\mathcal{G}f)(s, x) \right| \\ & \leq \frac{1}{t} \int_{s-t}^s |(G(s, r)\mathcal{A}(r)f(s-t, \cdot))(x) - (\mathcal{A}(s)f(s, \cdot))(x)| dr \\ & \quad + \frac{1}{t} \int_{s-t}^s |f_t(r, x) - f_t(s, x)| dr, \end{aligned}$$

for any $(s, x) \in \mathbb{R}^{d+1}$ and any $t > 0$. Arguing as in the proof of Proposition 6.1 it is immediate to check that the function $(r, p) \mapsto (G(s, r)\mathcal{A}(r)f(p, \cdot))(x)$ is continuous in $\{(r, p) \in \mathbb{R}^2 : r \leq s\}$. Therefore,

$$\lim_{t \rightarrow 0^+} (G(s, r)\mathcal{A}(r)f(s-t, \cdot))(x) - (\mathcal{A}(s)f(s, \cdot))(x) = 0,$$

for any $(s, x) \in \mathbb{R}^{d+1}$. Thus,

$$\lim_{t \rightarrow 0^+} \frac{(\mathcal{T}(t)f)(s, x) - f(s, x)}{t} = (\mathcal{G}f)(s, x), \quad (s, x) \in \mathbb{R}^{d+1}.$$

Moreover,

$$\begin{aligned} & \sup_{(s,x) \in [a,b+1] \times B_\delta} \left| \frac{(\mathcal{T}(t)f)(s, x) - f(s, x)}{t} \right| \\ & \leq \sup_{(r,s,x) \in [a,b+1] \times \text{supp}(\varphi)} |(\mathcal{A}(r)f(s, \cdot))(x)| + \|D_t f\|_\infty. \end{aligned}$$

Hence, the dominated convergence theorem implies that $t^{-1}(\mathcal{T}(t)f - f)$ converges to $\mathcal{G}f$, as $t \rightarrow 0^+$, in $L^p(\mathbb{R}^{d+1}, \nu)$ for any $p \in [1, +\infty)$. \square

7. AN EXAMPLE

In this section we consider operators $\mathcal{A}(t)$ defined on smooth functions $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$(7.1) \quad (\mathcal{A}(t)\varphi)(x) = \Delta\varphi(x) + \langle b(t, x), \nabla\varphi(x) \rangle,$$

under the following assumptions on $b = (b_1, \dots, b_d)$.

- Hypothesis 7.1.**
- (i) the functions b_j ($j = 1, \dots, d$) and their first-order spatial derivatives belong to $C_{loc}^{\frac{\alpha}{2}, \alpha}(I \times \mathbb{R}^d)$ for some $\alpha \in (0, 1)$;
 - (ii) the function $b(\cdot, 0)$ is bounded in I ;
 - (iii) there exists a continuous function $C : I \rightarrow \mathbb{R}$ such that
 - (a) C is bounded from above in I ;
 - (b) $\limsup_{t \rightarrow +\infty} C(t) < 0$;
 - (c) $\langle \nabla_x b(t, x)\xi, \xi \rangle \leq C(t)|\xi|^2$, $t \in I, x, \xi \in \mathbb{R}^d$.

Under Hypothesis 7.1, it is easy to check that, for any $N \in \mathbb{N}$, the function $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$, defined by $\varphi(x) = 1 + |x|^{2N}$ for any $x \in \mathbb{R}^d$, is a suitable Lyapunov function for the operator \mathcal{A} satisfying both Hypotheses 1.1(iii) and 1.4. Indeed, a straightforward computation shows that

$$(\mathcal{A}(t)\varphi)(x) = |x|^{2N-2} (4N^2 + 2N(d-2) + 2N\langle b(t, x), x \rangle).$$

Using Hypothesis 7.1(iii)(c), yields

$$b_j(t, x) = b_j(t, 0) + \int_0^1 \frac{d}{ds} b_j(t, sx) ds = b_j(t, 0) + \int_0^1 \langle \nabla_x b_j(t, sx), x \rangle ds,$$

so that

$$\begin{aligned} 2\langle b(t, x), x \rangle &= 2\langle b(t, 0), x \rangle + 2 \int_0^1 \langle \nabla_x b(t, sx)x, x \rangle ds \\ &\leq 2|b(t, 0)||x| + 2C(t)|x|^2, \end{aligned}$$

for any $t \in I$ and any $x \in \mathbb{R}^d$. Hence, for any $\varepsilon > 0$, we have

$$(7.2) \quad (\mathcal{A}(t)\varphi)(x) \leq (4N^2 + 2N(d-2))|x|^{2N-2} + 2|b(t, 0)||x|^{2N-1} + 2NC(t)|x|^{2N}.$$

Since

$$|x|^{2N-j} \leq \varepsilon|x|^{2N} + C_{\frac{2N}{j}}\varepsilon^{1-\frac{2N}{j}}, \quad x \in \mathbb{R}^N, \quad \varepsilon > 0, \quad j = 1, 2,$$

where $C_m = (m/(m-1))^{1-m}/m$, we can rewrite (7.2) as follows:

$$\begin{aligned} (A(t)\varphi)(x) &\leq \{2NC(t) + \varepsilon(2N(d-2) + 4N^2) + 2\varepsilon N|b(t, 0)|\}|x|^{2N} \\ &\quad + C_N(2N(d-2) + 4N^2)\varepsilon^{1-N} + 2NC_{2N}\varepsilon^{1-2N}|b(t, 0)| \\ (7.3) \quad &:= \psi_1(t)|x|^{2N} + \psi_2(t). \end{aligned}$$

Hypothesis 1.1(iii) follows taking $\varepsilon = 1$ and, for any bounded interval J compactly supported in I , $\lambda_J \geq \max\{\sup_J \psi_1, \sup_J \psi_2\}$. Similarly, if we fix $\varepsilon = \varepsilon_N$ such that

$$\varepsilon(d-2 + 2N + |b(\cdot, 0)|_\infty) \leq -\frac{1}{2} \limsup_{t \rightarrow +\infty} C(t),$$

and $t_0 \in \mathbb{R}$ such that

$$(7.4) \quad C(t) < \frac{1}{2} \limsup_{\tau \rightarrow +\infty} C(\tau), \quad t \geq t_0,$$

then Hypothesis 1.4 is satisfied for any $t \geq t_0$.

Finally, Hypothesis 1.3(ii) is trivially satisfied by virtue of Hypothesis 7.1(iii)(a). Hence, the following result follows.

Theorem 7.2. *Let \mathcal{A} be defined by (7.1) with the function b satisfying Hypothesis 7.1. Then, problem (1.2) is well posed in $C_b(\mathbb{R}^d)$. The corresponding evolution family $\{G(t, s)\}$ is irreducible and maps bounded measurable functions into bounded continuous functions. Moreover, $\{G(t, s)\}$ admits an evolution system (μ_t) of measures having bounded moments of any order $N \in \mathbb{N}$ for any $t \geq t_0$, where t_0 is the number in (7.4). In particular, all the polynomial functions are integrable with respect to μ_t for any $t \geq t_0$.*

REFERENCES

- [1] P. ACQUISTAPACE: *Evolution operators and strong solutions of abstract linear parabolic equations*, Diff. Int. Eqns. **1** (1988), pp. 433-457.
- [2] S. BERNSTEIN: *Sur la généralisation du problème de Dirichlet, I*, Math. Ann. **62** (1906), pp. 253-271.
- [3] M. BERTOLDI, L. LORENZI: *Analytical Methods for Markov Semigroups*, Pure and applied mathematics **283**, Chapman Hall/CRC Press, 2006.
- [4] C. CHICONE, Y. LATUSHKIN: *Evolution Semigroups in Dynamical Systems and Differential Equations*, Math. Surv. Monogr. **70**, AMS, Providence RI, 1999.
- [5] G. DA PRATO, M. RÖCKNER: *A note on non autonomous stochastic differential equations*. Proceedings of the 5th Seminar on Stochastic Analysis, Random Fields and Applications, Ascona 2005, R. Dalang, M. Dozzi, F. Russo Eds. Progress in Probability, Birkhäuser, (to appear).
- [6] G. DA PRATO, A. LUNARDI: *Ornstein-Uhlenbeck operators with time periodic coefficients*, J. Evol. Equ. **7** (2007), pp. 587-614.
- [7] E.B. DYNKIN: *Markov processes*, Grundlehren der Mathematischen Wissenschaften **121**, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1965.
- [8] S. FORNARO, G. METAFUNE, E. PRIOLA: *Gradient estimates for parabolic Dirichlet problems in unbounded domains*, J. Differential Equations **205**(2) (2004), pp. 329-353.
- [9] A. FRIEDMAN: *Partial differential equations of parabolic type*, Prentice Hall, 1964.
- [10] M. GEISSERT, A. LUNARDI: *Invariant measures and maximal L^2 regularity for nonautonomous Ornstein-Uhlenbeck equations*, (submitted).
- [11] M. GEISSERT, A. LUNARDI: *Asymptotic behavior in nonautonomous Ornstein-Uhlenbeck equations*, (submitted).
- [12] I. KARATZAS, S.E. SHREVE: *Brownian motion and stochastic calculus*. Second edition. Graduate Texts in Mathematics **113**. Springer-Verlag, New York, 1991.
- [13] H. KUNITA: *Stochastic flows and stochastic differential equations*. Reprint of the 1990 original. Cambridge Studies in Advanced Mathematics **24**, Cambridge University Press, Cambridge, 1997.
- [14] N. IKEDA, S. WATANABE: *Stochastic differential equations and diffusion processes*. Second edition. North-Holland Mathematical Library **24**, North-Holland Publishing Co., Amsterdam, 1989.
- [15] O.U. LADYŽENSKAJA, V.A. SOLONNIKOV, N.N. URAL'CEVA: *Linear and quasilinear equations of parabolic type*, Nauka, Moskow 1967 (Russian). English transl.: American Mathematical Society, Providence, 1968.
- [16] G.M. LIEBERMAN: *Second order parabolic differential equations*, World Scientific Publishing Co. Inc., River Edge, N.J., 1996.
- [17] G. METAFUNE, D. PALLARA, M. WACKER: *Feller Semigroups on \mathbb{R}^N* , Semigroup Forum **65**(2) (2002), pp. 159-205.

MARKUS KUNZE: GRADUIERENKOLLEG 1100 ULM UNIVERSITY
 HELMHOLTZSTRASSE 18, 89069 ULM, GERMANY
E-mail address: markus.kunze@uni-ulm.de

L. LORENZI, A. LUNARDI: DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DEGLI STUDI DI PARMA,
 VIALE G.P. USBERTI, 53/A, I-43100 PARMA, ITALY
E-mail address: luca.lorenzi@unipr.it
E-mail address: alessandra.lunardi@unipr.it